

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:54:12 ; Search time 2718 Seconds

(without alignments)
14904.748 Million cell updates/sec

Title: US-09-816-467-3
Sequence: 1392

1 atgggtttttcaacacccaat.....aggagatgacaaatgatataa 1392

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1387.4	99.7	1791	1	CLOTETTOX	M12739 C.tetani te
2	1387.4	99.7	1851	6	E01135	E01135 CDNA encodi
3	1387.4	99.7	4338	1	CTTETTOXG	X06214 Clostridium
4	1387.4	99.7	4400	1	CTTETANG	X04436 Clostridium
5	1379.8	99.1	1858	6	AR017561	AR017561 Sequence
6	1350.4	97.0	3931	1	AF389424	AF389424 Clostridi
7	1346.4	96.7	1359	6	A17527	A17527 human fragm
8	1346.4	96.7	1359	6	I28430	I28430 Sequence 1
9	1336.8	96.0	1356	1	AF154828	AF154828 Clostridi
10	1287.2	92.5	1359	6	I13981	I13981 Sequence 7
11	799.2	57.4	1359	6	I28431	I28431 Sequence 3
12	799.2	57.4	1359	6	I28431	I28431 Sequence 3
13	799.2	57.4	1359	6	I28431	I28431 Sequence 3
14	797.2	57.3	4366	6	A49987	A49987 Sequence 4
15	796.8	57.2	3754	6	A42484	A42484 Sequence 12
16	796.8	57.2	3754	6	A37074	A37074 Sequence 17
17	796.8	57.2	3769	6	A42478	A42478 Sequence 6
18	796.8	57.2	3769	6	A49988	A49988 Sequence 18
19	796.8	57.2	4378	6	A42481	A42481 Sequence 5
20	521	37.4	1766	6	A49989	A49989 Sequence 9
21	248.2	17.8	3937	1	CBBONTG	X74162 C.botulinum
22	236.4	17.0	3876	1	AF295926	AF295926 Clostridi
23	235.8	16.9	4041	1	CLOBOTB	M81166 Clostridium
24	235.8	16.9	11170	1	CBY13630	Y13630 Clostridium
25	234.8	16.9	3876	1	AF300465	AF300465 Clostridi
26	234.8	16.9	3876	1	AF300465	AF300465 Clostridi
27	233.8	16.8	4051	1	CBBOBNT	X71343 C.botulinum
28	232	16.7	1326	1	CBBO242628	AR000031 Sequence
29	228.2	16.4	3891	6	AR000031	AR000031 Sequence
30	228.2	16.4	3891	6	AR000031	AR000031 Sequence
31	228.2	16.4	3891	6	AR000031	AR000031 Sequence
32	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
33	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
34	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
35	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
36	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
37	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
38	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
39	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
40	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
41	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
42	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
43	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
44	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence
45	228.2	16.4	4292	1	CBBO242628	AR000031 Sequence

ALIGNMENTS

RESULT 1
LOCUS CLOTETTOX
DEFINITION C.tetani tetanus toxin gene.
ACCESSION M12739.1 GI:144920
VERSION M12739.1 GI:144920
KEYWORDS tetanus toxin.
SOURCE C.tetani (strain CN3911, derivative of Harvard strain) DNA, clones phtell, 81.
ORGANISM Clostridium tetani
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1 (bases 1 to 1791)
Fairweather,N.F., Lyness,V.A., Pickard,D.J., Allen,G. and
Thomson,R.O.

Pred. No. is the number of results predicted by chance to have a

TITLE Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in *Escherichia coli*
 JOURNAL J. Bacteriol. 165 (1), 21-27 (1986)
 MEDLINE 86085672
 PUBMED 3510187
 COMMENT Draft entry and sequence in computer readable form for [1] kindly provided by N.F. Fairweather, 05-AUG-1986.
 the last 122 amino acids of fragment B are found at position 1 to 366. Fragment C of the tetanus toxin is located at positions 367-1719.

FEATURES

source

1. 1791
 /organism="Clostridium tetani"
 /db_xref="taxon:1513"
 <1..1722
 /note="tetanus toxin"
 /codon_start=1
 /transl_table=1
 /protein_id="AA023282.1"
 /db_xref="GI:144921"

CDS

/translation="RSLEYOVDAIKRLIDYEYKISGPRDEQIADENINLKNLEEKANAMINININRESRSFLVNMINEAKKOLLEEDTOSKINLMQYIKANSKFLITELKLESKINKEVSPGIPESYKNDLQWVDNEDIDILKSTLILNDIINDIISDISEFSSSVITYPDQOLVPEINGKAIHLVNNSESVIYHKAMDIEYNDENNPTESEMLRVPKVSASHELOYGTNEYSIISMKHSLSTGSGSVSLKGNLITWTKDSAGEVROITPDLDPKFNAYLANRWFITITNDRLSSANDLYNGVLMGSEITGLGALREDNNTLKLRCSNNNOYVSIIDKRFICKALNPKEIKELYSITLPELRODGNPLRDEYLIPIVASSKDOVKNTITDYMLNAPSYNGKNIYSLRKYNGKRIKRYRPNNDIIDEFVSGDFIKLYVNNNHIVGYPKDGNAFNNDRLILVGYNAPKIPLYKMEAVKLDLKTYSVOLKLYDXKNASLGLVGTINQIGNDPNNDILIASWVYNHLKDKILGCDWYFVPTDEGMTND

BASE COUNT 727 a 197 c 276 g 591 t
 ORIGIN 1 bp upstream of BglII site.

Query Match 99.7% Score 1387.4; DB 1; Length 1791;
 Best Local Similarity 99.9%; Pred. No. 3.2e-174;

Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 GTTTTTCACCAATTCATTTCTTATCTTAAATCTGATTTGGTGTGATTAAT 63
 DB 334 GTTTTTCACCAATTCATTTCTTATCTTAAATCTGATTTGGTGTGATTAAT 393
 OY 64 GAAGAATATGATGTATTAATAAAGAGTACATTTTAATTTAGATTTTAATAT 123
 DB 394 GAAGAATATGATGTATTAATAAAGAGTACATTTTAATTTAGATTTTAATAT 453
 OY 124 GATATTATATCAGATATATCTGGTTAATTCATCTGTANTATCAATCCAGTCTCAA 183
 DB 454 GATATTATATCAGATATATCTGGTTAATTCATCTGTANTATCAATCCAGTCTCAA 513
 OY 184 TTTGGTCCCGGATTAATGGCAAGCAATACATTTAGTAACAATGCTTGAAGTT 243
 DB 514 TTTGGTCCCGGATTAATGGCAAGCAATACATTTAGTAACAATGCTTGAAGTT 573
 OY 244 ATAGTGCATTAAGCTATGATTTGAATATTAAGATATGTTTAATTTTACCGTTAGC 303
 DB 574 ATAGTGCATTAAGCTATGATTTGAATATTAAGATATGTTTAATTTTACCGTTAGC 633
 OY 304 TTTTGGTTGAGGGTCTCTAAAGTATCTGCTAGTCATTTAACAATTTGGCACAAATGAG 363
 DB 634 TTTTGGTTGAGGGTCTCTAAAGTATCTGCTAGTCATTTAACAATTTGGCACAAATGAG 693
 OY 364 TATTCATATTTAGCTCTATGAAAAAACAATAGCTATCATAGATAGGATCGTTGAGTGA 423
 DB 694 TATTCATATTTAGCTCTATGAAAAAACAATAGCTATCATAGATAGGATCGTTGAGTGA 753
 OY 424 TCACCTTAAGGTAATACTTAATATGAGCTTTAAAGATTTCCGCGGAGAGATTAGCAAA 483
 DB 754 TCACCTTAAGGTAATACTTAATATGAGCTTTAAAGATTTCCGCGGAGAGATTAGCAAA 813
 OY 484 ATAACTTTTAGGGATTACCGATTAATTTTAATGCTTATTAGCAAAATTAATGGGTTTTT 543
 DB 814 ATAACTTTTAGGGATTACCGATTAATTTTAATGCTTATTAGCAAAATTAATGGGTTTTT 873

OY 544 ATAACTTTTAGGGATTACCGATTAATTTTAATGCTTATTAGCAAAATTAATGGGTTTTT 603
 DB 874 ATAACTTTTAGGGATTACCGATTAATTTTAATGCTTATTAGCAAAATTAATGGGTTTTT 933
 OY 604 GGAAGTCAGAAATTTAGCTGTTTGGAGCTATTAGAGAGATAATATTAATACATTAAAA 663
 DB 934 GGAAGTCAGAAATTTAGCTGTTTGGAGCTATTAGAGAGATAATATTAATACATTAAAA 993
 OY 664 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
 DB 994 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1053
 OY 724 AAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
 DB 1054 AAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1113
 OY 784 TTAAGAGACTTCTGGGGAACCCCTTACGATATGATACAGATATTAATTAATTAATTAAT 843
 DB 1114 TTAAGAGACTTCTGGGGAACCCCTTACGATATGATACAGATATTAATTAATTAATTAAT 1173
 OY 844 GCTTCTAGTCTTAAGATGTTCAATTTGAATAATTAATTAATTAATTAATTAATTAAT 903
 DB 1174 GCTTCTAGTCTTAAGATGTTCAATTTGAATAATTAATTAATTAATTAATTAATTAAT 1233
 OY 904 GCGGCATGCTATACAAAGAAATTTGAATATTAATTAATTAATTAATTAATTAATTAAT 963
 DB 1234 GCGGCATGCTATACAAAGAAATTTGAATATTAATTAATTAATTAATTAATTAATTAAT 1293
 OY 964 AAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
 DB 1294 AAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1353
 OY 1024 GATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
 DB 1354 GATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1413
 OY 1084 GATGGAATGCTTAAATATATCTTGTAGAAATTTCTAAGATAGGTTAATAGCCCAAGGT 1143
 DB 1414 GATGGAATGCTTAAATATATCTTGTAGAAATTTCTAAGATAGGTTAATAGCCCAAGGT 1473
 OY 1144 ATCCCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
 DB 1474 ATCCCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1533
 OY 1204 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
 DB 1534 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593
 OY 1264 CAATATGCGCAAGATCCAAATAGGATATTTAATTCGAAGCAACTGCTTAATCAT 1323
 DB 1594 CAATATGCGCAAGATCCAAATAGGATATTTAATTCGAAGCAACTGCTTAATCAT 1653
 OY 1324 TTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1383
 DB 1654 TTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
 OY 1384 AATGATTAA 1392
 DB 1714 AATGATTAA 1722

RESULT 2
 E01135 1851 bp DNA linear PAT 29-SEP-1997
 LOCUS E01135
 DEFINITION cDNA encoding tetanus toxin C fragment.
 ACCESSION E01135
 VERSION E01135.1 GI:2169394
 KEYWORDS JP 1987051994-A/1.
 SOURCE Clostridium tetani.
 ORGANISM Clostridium tetani.
 Clostridia; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

REFERENCE 1 (bases 1 to 1851)
AUTHORS Nei, R. F.
TITLE CLONED ANTIGEN
JOURNAL Patent: JP 1987051994-A 1 06-MAR-1987;
WELLCOME FOUND LTD:THE
COMMENT OS Clonitridium tetani
PN JP 1987051994-A/1
PD 06-MAR-1987
PF 27-JUN-1986 JP 1986149849
PR 28-JUN-1985 GB 85 8516442
PI NEIRU FUREIZAA FUEAUEZAA
PC C12N15/00, A61K39/08, C07K13/00, C12P21/02, (C12N15/00, C12R1:145),
PC (C12P21/02,
PC C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain-CN 3911;
FH key Location/Qualifiers
FT CDS
FT FT
FT mat_peptide
FT /product='tetanustoxin'
FT 3'UTR
FT mRNA
FT Location/Qualifiers
1.1851
/organism='Clostridium tetani'
/db_xref='taxon:1513'
BASE COUNT 753 a 204 c 279 g 615 t
ORIGIN
Query Match 99.7%; Score 1387.4; DB 6; Length 1851;
Best Local Similarity 99.9%; Pred. No. 3.2e-174;
Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GTTTTTCACACCAATTCATTTTCTTATTAATAAATCTGGATTTGGGTGATAT 63
DB 334 GTTTTTCACACCAATTCATTTTCTTATTAATAAATCTGGATTTGGGTGATAT 393
QY 64 GAAGAAGATTTAGTGTATTAATAAAGAGTACATTTTAATTTAGATTAATAT 123
DB 394 GAAGAAGATTTAGTGTATTAATAAAGAGTACATTTTAATTTAGATTAATAT 453
QY 124 GATATTAATACATATATCTGGTTTAATTCATCTGTAATACATATCCAGATGCTCA 183
DB 454 GATATTAATACATATATCTGGTTTAATTCATCTGTAATACATATCCAGATGCTCA 513
QY 184 TTGGTCCCGGATTAATGCAAGCAATACATTTAGTAAACAATGAATCTTGAAGTT 243
DB 514 TTGGTCCCGGATTAATGCAAGCAATACATTTAGTAAACAATGAATCTTGAAGTT 573
QY 244 ATAGTGCATTAACCTATGATTTGATTAATGATATGTTTAAATTTTACCGTTAGC 303
DB 574 ATAGTGCATTAACCTATGATTTGATTAATGATATGTTTAAATTTTACCGTTAGC 633
QY 304 TTTTGGTGGAGGTTCCATAAGATCTGCTAGTCAATTTAGAACATATGGCACAAATGAG 363
DB 634 TTTTGGTGGAGGTTCCATAAGATCTGCTAGTCAATTTAGAACATATGGCACAAATGAG 693
QY 364 TATTCATATTAATGCTATGAAAAACATAGCTATCATATAGGATTTGGTGGAGTGA 423
DB 694 TATTCATATTAATGCTATGAAAAACATAGCTATCATATAGGATTTGGTGGAGTGA 753
QY 424 TCACCTAAAGGTAATTAATTAATGAGCTTTAAAGATTTCCCGGAGAGAGTTAGACAA 483
DB 754 TCACCTAAAGGTAATTAATTAATGAGCTTTAAAGATTTCCCGGAGAGAGTTAGACAA 813
QY 484 ATAACTTTAGGATTTACCTGATAAATTAATGCTTATTTAGCAATAAATGGGTTT 543

DB 814 ATAACTTTAGGATTTACCTGATAAATTAATGCTTATTTAGCAATAAATGGGTTT 873
QY 544 ATAACTTTAGGATTTACCTGATAAATTAATGCTTATTTAGCAATAAATGGGTTT 603
DB 874 ATAACTTTAGGATTTACCTGATAAATTAATGCTTATTTAGCAATAAATGGGTTT 933
QY 604 GGAAGTCAGAAATTTACGTTTGGAGCTTTAGAGAGGTAATTAATTAATTAATTAAT 663
DB 934 GGAAGTCAGAAATTTACGTTTGGAGCTTTAGAGAGGTAATTAATTAATTAATTAAT 993
QY 664 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
DB 994 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1053
QY 724 AAAGCATTAATCCAAAAGAGATTTGAATTAATTAATTAATTAATTAATTAATTAAT 783
DB 1054 AAAGCATTAATCCAAAAGAGATTTGAATTAATTAATTAATTAATTAATTAATTAAT 1113
QY 784 TTAAGAGACTTTGGGGAACCCCTTTAGATATGATACAGAAATTAATTAATTAATTAAT 843
DB 1114 TTAAGAGACTTTGGGGAACCCCTTTAGATATGATACAGAAATTAATTAATTAATTAAT 1173
QY 844 GCTTCTAGTCTTAAGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
DB 1174 GCTTCTAGTCTTAAGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1233
QY 904 GGGCCATCGTATCTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
DB 1234 GGGCCATCGTATCTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1293
QY 964 AAATTTATTAATAAAGATTTATACCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
DB 1294 AAATTTATTAATAAAGATTTATACCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1353
QY 1024 GATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
DB 1354 GATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1413
QY 1084 GATGGAATGCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
DB 1414 GATGGAATGCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1473
QY 1144 ATCCCTCTTTATTAATAAAGATTTGAAGCAATTAATTAATTAATTAATTAATTAATTAAT 1203
DB 1474 ATCCCTCTTTATTAATAAAGATTTGAAGCAATTAATTAATTAATTAATTAATTAATTAAT 1533
QY 1204 CAACCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
DB 1534 CAACCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593
QY 1264 CAATATAGGCAAGATCAATATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
DB 1594 CAATATAGGCAAGATCAATATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1633
QY 1324 TTAAGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1383
DB 1634 TTAAGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
QY 1384 AATGATTA 1392
DB 1714 AATGATTA 1722
RESULT 3
CTTETXG 4338 bp DNA linear BCT 12-SEP-1993
LOCUS Clostridium tetani gene for tetanus toxin.
DEFINITION X06214
ACCESSION X06214.1 GI:40773
VERSION X06214.1
KEYWORDS tetanus toxin.
SOURCE Clostridium tetani.
ORGANISM Clostridium tetani.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

REFERENCE 1 (bases 1 to 4338)
 Fairweather, N.F. and Lyness, V.A.
 TITLE The complete nucleotide sequence of tetanus toxin
 JOURNAL Nucleic Acids Res. 14 (19), 7809-7812 (1986)
 MEDLINE 87040747
 PUBMED 3774547
 COMMENT Data kindly reviewed (25-Nov-1988) by Lyness V.A.
 FEATURES
 Location/Qualifiers
 1. 4338
 /organism="Clostridium tetani"
 /strain="CN3911"
 /db_xref="taxon:1513"
 /clone="PAT153"
 322..4269
 /note="tetanus toxin (AA 1-1315)"
 /codon_start=1
 /protein_id="CAA29564.1"
 /db_xref="GI:40774"
 /db_xref="SWISS-PROT:P04958"
 /translation="MPITINFRYSDPVNNDTIIMPEPYCKGLDIYKAFKITDRIM
 IYPEREFGKPEDNPSSLIIGASEYEDPNLFDSDKDFLOPMVKLFNIKNNV
 AGEALLDKIINAPYGNYSYLIDKEDTMSYSFMLFODPSGATTKSAMLNLIIF
 GREPVLNKNEVRCIVLRDNKNTFPCRDGSGSIMQAFCEPYEPEDNTENTISLT
 GSKRYQDPALMLHELIVLHGLYGMVSSHETPSKDEITMHTPIISAEELPFG
 GDQANLISIDKNDLYEKTLDYKALANKLSQVTSNDPDIIDISTYKQYQVFDK
 DSNQGYIVNEDKFOILYNSIMGFTEIELGKKFNITRSLYSFMNDHPVKIPMLDPT
 IYDTEGPFNIESDKSEYKQGMRYNATFRVDSGLVSKLIGLCKKIIPNTIRE
 NLYNRTASLIDGGEGLCKIKNEDLFTIEAKNFSEPFODEIVSNITKPKPLNFNS
 LDKITLYDNLSQKTIPLPNDTPPTKGIQIPAPYKNAASTIEIHNIINDIYQYLA
 QKSPQTLQRTITMNSVDLALINSTKIYSTRPYSIKVNOAGQITLLOWRDIIDPT
 NNESSQRTYIDKISDVSTIVYIGPALNIVQSGEGNIGALETTGVLLLEYTPETTL
 PVTAALSIESTSTOKKIKITIDNPLEKREKMEYKLVKAKWALGVNTQFQRSTQ
 MRSLEQVDAIKKIIDEYK IYSGPDEQIADENILNKKLEKANKAMININIMR
 ESSRSFLYNOMINEAKKOLEPDTOSKILMOYIKANSKFIGITELKLESKINKVES
 TPPIPSYSKRLDQVNEEDIDYLIKSTKILINDINNDISDISGESSVITYPDOL
 VPEINGKAIHLVNNSESEYIVHKAAMDIEYDMENNPVSMLEVPKVSASHLOYGN
 EYSIISSMKHSLISGMSVSLKGNLITLTLKDSAGEVQITPRDIPKFNAYLANK
 WVEITITNDRLSASNLINGVLMSALETGLAIREDNITLTLDRNNNOYVSDIK
 FRIFCKALNPKKEIKELVSYLSITFLDEFGNPLRDPEYLLIPVASSSDYOLKNT
 DYVLTNAPSYTNGKLNLYRRRLNGLKEFIKRYTPNNEIDSEVKSQDFIKLVSYNN
 NEHLYGPKDGNAPNLDRLIRYGVNAPGILPKYKMEVAILRDLKTYVOLKLYDKDN
 ASLIGVTHNGQIGNDPNDRIILIASNMYFHLMDKILGCMWYFVPTDEGMTND"

BASE COUNT 1783 a 522 c 640 g 1393 t
 ORIGIN

Query Match 99.7%; Score 1387.4; DB 1; Length 4338;
 Best Local Similarity 99.9%; Pred. No. 2.4e-174;
 Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTTTTTCACACCAATTCATTTCTTATTCCTAAATATCGGATGTTGGTGATAT 63
 DB 2881 GTTTTTCACACCAATTCATTTCTTATTCCTAAATATCGGATGTTGGTGATAT 2940
 QY 64 GAAGAAGTATAGATATATTAATAAAGAGTACATTTTAAATTTGATATTAAT 123
 DB 2941 GAAGAAGTATAGATATATTAATAAAGAGTACATTTTAAATTTGATATTAAT 3000
 QY 124 GATATTATATCAGATATATCTGGGTTTAAATTCATCTGTAATAACATTCAGATGCTCAA 183
 DB 3001 GATATTATATCAGATATATCTGGGTTTAAATTCATCTGTAATAACATTCAGATGCTCAA 3060
 QY 184 TTGGTGCCCGGAATAAAGCAACATACCTTTAGTAACAATAAGATCTTCGAAGTT 243
 DB 3061 TTGGTGCCCGGAATAAAGCAACATACCTTTAGTAACAATAAGATCTTCGAAGTT 3120
 QY 244 ATAGTGCATTAAGTATGATATTAATATAATGATATGATATTAATTTACCGTTAGC 303
 DB 3121 ATAGTGCATTAAGTATGATATTAATATAATGATATGATATTAATTTACCGTTAGC 3180
 QY 304 TTTTGGTGAGGGTTCCTAAGTATCTGCTAGTCATTTAGAACATATGGCAATGAG 363
 |||||||

DB 3181 TTTTGGTGAGGGTTCCTAAGTATCTGCTAGTCATTTAGAACAAATATGGCAAAATGAG 3240
 QY 364 TATTCATATATTAAGTCTCTATGAAGAAAAACATAGTCTATCATAGGATCTGGTGAGTGA 423
 |||||||
 DB 3241 TATTCATATATTAAGTCTCTATGAAGAAAAACATAGTCTATCATAGGATCTGGTGAGTGA 3300
 QY 424 TCACCTTAAGGATATATACCTTAATATGAGACTTTAAAGATCCCGGAGAGTATAGCAA 483
 |||||||
 DB 3301 TCACCTTAAGGATATATACCTTAATATGAGACTTTAAAGATCCCGGAGAGTATAGCAA 3360
 QY 484 ATACCTTTTAGGATTTTACCTGATTAATATTAATGCTTATTTAGCAATAAATGGGTTTTT 543
 |||||||
 DB 3361 ATACCTTTTAGGATTTTACCTGATTAATATTAATGCTTATTTAGCAATAAATGGGTTTTT 3420
 QY 544 ATACCTTTTACTATGATATATATATATCTTCGCTAATTTGATATATAATGAGATCTATG 603
 |||||||
 DB 3421 ATACCTTTTACTATGATATATATATATCTTCGCTAATTTGATATATAATGAGATCTATG 3480
 QY 604 GGAAGTGCAGAAATTTACTGGTTTAGAGACTATTAGAGAGTAAATTAATTAACATTAATA 663
 |||||||
 DB 3481 GGAAGTGCAGAAATTTACTGGTTTAGAGACTATTAGAGAGTAAATTAATTAACATTAATA 3540
 QY 664 CTAGATAGATGTAATTAATTAATTAATCAATPAGCTTCTATGATAATTTAGATATTTTGC 723
 |||||||
 DB 3541 CTAGATAGATGTAATTAATTAATTAATCAATPAGCTTCTATGATAATTTAGATATTTTGC 3600
 QY 724 AAGCATTAATTCAGAAAGATTTGAAAGAAATTTATACCAAGTTATTAATCTATACCTTT 783
 |||||||
 DB 3601 AAGCATTAATTCAGAAAGATTTGAAAGAAATTTATACCAAGTTATTAATCTATACCTTT 3660
 QY 784 TTAAGACACTTCTGGGGAACCCCTTTACGATATGATACAGAAATTTATTAATACAGTA 843
 |||||||
 DB 3661 TTAAGACACTTCTGGGGAACCCCTTTACGATATGATACAGAAATTTATTAATACAGTA 3720
 QY 844 GCTTCTAGTCTTAAGATGTTCAATTTGAAAATATTAACAGATTTATGATTTTGGCAAT 903
 |||||||
 DB 3721 GCTTCTAGTCTTAAGATGTTCAATTTGAAAATATTAACAGATTTATGATTTTGGCAAT 3780
 QY 904 GCGGCATGCTTACTTAACGAGAAATTTGAATATTAATTAAGAGTTTATTAATGAGACTA 963
 |||||||
 DB 3781 GCGGCATGCTTACTTAACGAGAAATTTGAATATTAATTAAGAGTTTATTAATGAGACTA 3840
 QY 964 AAATTTATTAAGAAAGATTAACCTTAATTAATGAATAGATTTCTTTGTTAAATCAGGT 1023
 |||||||
 DB 3841 AAATTTATTAAGAAAGATTAACCTTAATTAATGAATAGATTTCTTTGTTAAATCAGGT 3900
 QY 1024 GATTTTATTAATTAATTAATGATATATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
 |||||||
 DB 3901 GATTTTATTAATTAATTAATGATATATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
 QY 1084 GATGAAGATGCTTAAATTAATCTTGATAGAAATTTCTAAGAGTAGTATTAATGCCAGGT 1143
 |||||||
 DB 3961 GATGAAGATGCTTAAATTAATCTTGATAGAAATTTCTAAGAGTAGTATTAATGCCAGGT 4020
 QY 1144 ATCCCTCTTTATTAAGAAATTAAGACAGTAAATTTGCGTGATTTAAAGAACTATCTGTA 1203
 |||||||
 DB 4021 ATCCCTCTTTATTAAGAAATTAAGACAGTAAATTTGCGTGATTTAAAGAACTATCTGTA 4080
 QY 1204 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
 |||||||
 DB 4081 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
 QY 1264 CAATATAGCAACGATCCAAATAGGATATATTAATTAATTAATTAATTAATTAATTAAT 1323
 |||||||
 DB 4141 CAATATAGCAACGATCCAAATAGGATATATTAATTAATTAATTAATTAATTAATTAAT 4200
 QY 1324 TTTAAAGATTAATTTTATAGAGTATGATGCTTGTACCTACAGATAGAGGATGAGACA 1383
 |||||||
 DB 4201 TTTAAAGATTAATTTTATAGAGTATGATGCTTGTACCTACAGATAGAGGATGAGACA 4260
 QY 1384 AATGATTAAT 1392
 |||||||
 DB 4261 AATGATTAAT 4269

RESULT 4
LOCUS CTETANG 4400 bp DNA linear BCT 12-JUL-1995
DEFINITION Clostridium tetani gene for tetanus toxin.
ACCESSION X04436
VERSION X04436.1 GI:40769
KEYWORDS neurotoxin; plasmid; tetanus toxin; toxin.
ORGANISM Clostridium tetani.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE 1 (bases 1 to 4400)
AUTHORS Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hidel, M., Habermann, E. and Niemann, H.
TITLE Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxins
JOURNAL EMBO J 5 (10), 2495-2502 (1986)
MEDLINE 87053814
PubMed 3536478
COMMENT See also Fairweather et al. (1986) J. Bacteriol. 165, 21-27. Data kindly reviewed (12-MAY-1987) by U. Eisel.
FEATURES
source location/Qualifiers
1..4400
/organism="Clostridium tetani"
/db_xref="taxon:1513"
<1..50
/note="Open reading frame (16AA) (1 is 2nd base in codon)"
/codon_start=3
/transl_table=1
/protein_id="CAA28032.1"
/db_xref="GI:809667"
/db_xref="SPTREMBL:Q46456"
/translation="ALKLEPIVVKLNI"
37..250
/note="region of inverted repeats"
267..272
/note="rRNA binding site"
281..4228
/codon_start=1
/transl_table=1
/product="tetanus toxin precursor (AA 1-1315)"
/protein_id="CAA28033.1"
/db_xref="GI:40780"
/db_xref="SWISS-PROT:P04958"
/translation="MPTINNFRYSDPVNNDIIMMEPYCKGLDIYYKAEKITDRIM IVERYEFGRKPEDFNPSSLIEGASEYDNYLTDSKDRFLOTWKLEFRIKNY AGEALDKIINALPYLGNYSILDKPFPNSVSFNLEOPDSGATTSAMITNLITF GPGVFLKNKEVRGIVLRVKNKPYPCBDFGCSIMOMAFCPYVPEFDVETITSLTI GKSKYFDDPALMLHELIHYLHGIXGMQVSHETIPSKOETIMORTYPISEELTFG GQANLISIDIKNDLYEKTLLNDYKAIANKLSQVTSQNDPIDDSYKQIYQOKYQDK DSKQYIVNEDKRFQILNYSIMYGFTELEGKFNKRLSYFSNHDVVKIPLNDPDT IYNDTEGFNIESDKSEYKQNMKNVNFARNVDSGLVGLGCKIIPNLIRE NLNRTASLTDLGEGLCKIKINEDLTFIAEKNSFSEFOEIEIVSNKKNPNINS LDKIIVYNLQSKITLNDRTTPYTKGIPVAPKESNAQSTIEIHNDNNIYQVLYA OKSPPTORTITMNSVDALINSTKIYSPYSKYNQAGLIFLOWARDITDDFT NISSOKTIDKISVSTIVYIGPALNIVKQYSGNFGALETTGVLLEIPEITL PYTAASIASSSTOKETIKITIDNLEKREKLEVKLVAKMGLVNTQFORSYO MYRSLEYOVAIKRIIDYEKIVSGPDKLEQIADENINKNLEKANKAMININEMR ESSRSLVNMQINEAKROLEPDSQNIIMQYIKANSFGITELKLESKINVS TPIPFYSKNDICVNEEDIDVILKSKTILINDINNDISIDIGFNSVITYPAOL VPGINGKAIHLVNNSESEVIVKAMDIEYNDMPNFPYSPFLRPKVSASHLEOYGN EYSIISMKHSLSTIGSGMSVSLKGNLITGLKAGAGVROITREDLDFKNAVIANK KVIITINDLSANLITNGVLMGSABITGGAIRDNNTITLKDRCNNNNQVSIADK FRIEFCALNKEIEKLTYSYSTLFLDFWGNPNIKYDEYVLLIVASSSKDVQKNTI DYMLITNAPSYTNGKLNLYRRLNGLKFIIRYTPNNEIDSEVSKSGDFIKLYSYNN NEHIVGPKXGNAFNNIDRLIRVGNAPGIPLYKKMAVLRKDRDKTSVOLKIVDDKN ASLIGVTHNGQIGNDPNRIDILASNMVFHMKKIKIIGCDWYFVPTDEGTND"
281..1651
/product="tetanus toxin L-chain (AA 1-457)"
1652..4225
/product="tetanus toxin H-chain (AA 458-1315)"
mat_peptide
mat_peptide

misc_feature 4268..4394
/note="region of inverted repeats"
BASE COUNT 1808 a 532 c 644 g 1416 t
ORIGIN
Query Match 99.7%; Score 1387.4; DB 1; Length 4400;
Best Local Similarity 99.9%; Pred. No. 2.4e-174;
Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GTTITTTCAACACCAATTCATTTTCTTAAATAATCTGGATTGGGTGATAAT 63
|||||
Db 2840 GTTITTTCAACACCAATTCATTTTCTTAAATAATCTGGATTGGGTGATAAT 2899
QY 64 GAAGAGATATAGATATATATATATATATATATATATATATATATATATAT 123
|||||
Db 2900 GAAGAGATATAGATATATATATATATATATATATATATATATATATATAT 2959
QY 124 GATATATATAGATATATATATATATATATATATATATATATATATATAT 183
|||||
Db 2960 GATATATATAGATATATATATATATATATATATATATATATATATATAT 3019
QY 184 TTGGTGGCCGGAATTAATAGCAACATACATTAGTAACAAATGATCTTGAAGTT 243
|||||
Db 3020 TTGGTGGCCGGAATTAATAGCAACATACATTAGTAACAAATGATCTTGAAGTT 3079
QY 244 ATAGTGAATTAAGTATAGATATATATATATATATATATATATATATATAT 303
|||||
Db 3080 ATAGTGAATTAAGTATAGATATATATATATATATATATATATATATATAT 3139
QY 304 TTTTGGTTGAGGGTTCCTAAGATATCTGCTAGTCATTTAGAACAAATATGCA 363
|||||
Db 3140 TTTTGGTTGAGGGTTCCTAAGATATCTGCTAGTCATTTAGAACAAATATGCA 3199
QY 364 TATTCATTAATTTACCTCTATGAAAAAACAATAGTATGATAGATCTGGTGGAGTGA 423
|||||
Db 3200 TATTCATTAATTTACCTCTATGAAAAAACAATAGTATGATAGATCTGGTGGAGTGA 3259
QY 424 TCACCTTAAGGTATTAACCTATATATGACTTTAAAGATTCGCGGAGAACTTAGACA 483
|||||
Db 3260 TCACCTTAAGGTATTAACCTATATATGACTTTAAAGATTCGCGGAGAACTTAGACA 3319
QY 484 ATACCTTTTAGGGATTTACCTGATTAATTTAATGCTTATTTAGCAATTAATGGGTTTT 543
|||||
Db 3320 ATACCTTTTAGGGATTTACCTGATTAATTTAATGCTTATTTAGCAATTAATGGGTTTT 3379
QY 544 ATACCTTTACTAATGATATGATATATCTGCTAATTTGATATATATAGATCTTATAG 603
|||||
Db 3380 ATACCTTTACTAATGATATGATATATCTGCTAATTTGATATATATAGATCTTATAG 3439
QY 604 GGAAGTGCAGAAATTTACTGTTTAGAGAGCTATTAGAGAGATTAATATACATTAATA 663
|||||
Db 3440 GGAAGTGCAGAAATTTACTGTTTAGAGAGCTATTAGAGAGATTAATATACATTAATA 3499
QY 664 CTGATATGAT 723
|||||
Db 3500 CTGATATGAT 3559
QY 724 AAAGCATTAATCAAAAAGATGGAATAATTTATACACATTTATTTCTATACCTTT 783
|||||
Db 3560 AAAGCATTAATCAAAAAGATGGAATAATTTATACACATTTATTTCTATACCTTT 3619
QY 784 TTAAGACATTTGCGGGAACCTTTACGATATGATACAGAAATTTATTTATATACAGTA 843
|||||
Db 3620 TTAAGACATTTGCGGGAACCTTTACGATATGATACAGAAATTTATTTATATACAGTA 3679
QY 844 GCTTTAGATTCTAAAGATGTTCAATGAATAATTTATACAGATTTATGTTATGCAAAAT 903
|||||
Db 3680 GCTTTAGATTCTAAAGATGTTCAATGAATAATTTATACAGATTTATGTTATGCAAAAT 3739
QY 904 GCGGCATGCTATATACGGAATAATGAATATATATATATATATATATATATATATAT 963
|||||
Db 3740 GCGGCATGCTATATACGGAATAATGAATATATATATATATATATATATATATATATAT 3799

QY	964	AAATTTTATTTTAAAAAACAATATPACCCCAATAAAGAATAGATCTCTTTGGTTAAATCAGGT	1023
Db	3800	AAATTTTATTTTAAAAAAGATATPACACCTAATPATGAAATAGATCTCTTTGGTTAAATCAGGT	3859
QY	1024	GATTTTATTTTAAATATATATGATCATATATACAAATTAATGAGCAGATTGAGTTAATCCGAAA	1083
Db	3860	GATTTTATTTTAAATATATATNGTATCATATATACAAATTAATGAGCAGATTGAGTTATCCGAAA	3919
QY	1084	GATGGAATGCGCTTTAATPAATCTTGATAGAAATCTAAAGTAGTGTATATGCCCAAGGT	1143
Db	3920	GATGGAATGCGCTTTAATPAATCTTGATAGAAATCTTAAGAGAGGTATATAAGCCCCAGGT	3979
QY	1144	ATCCCTCTTTTATAAAAAATGGAAGCAGTAAATTTGGGTGATTTAAAAAACCCTATCTGTA	1203
Db	3980	ATCCCTCTTTTATAAAAAATGGAAGCAGTAAATTTGGGTGATTTAAAAAACCCTATCTGTA	4039
QY	1204	CAACTTAATATATATGATGATGATAAAAATGCAATCTTTAGGACTAGTAGTACCATAATGCT	1263
Db	4040	CAACTTAATATATATGATGATGATAAAAATGCAATCTTTAGGACTAGTAGTACCATAATGCT	4099
QY	1264	CAAAATGAGCAACGATCCCAATAGGAGTATATTAATTGCAAGCACTGCTACTTTAATCAT	1323
Db	4100	CAAAATGAGCAACGATCCCAATAGGAGTATATTAATTGCAAGCACTGCTACTTTAATCAT	4159
QY	1324	TTTAAAGTATAAATTTTATGAGTGTGATGTTGACTCTTGACTACAGATGAGAGATGAGCA	1383
Db	4160	TTTAAAGTATAAATTTTATGAGTGTGATGTTGACTCTTGACTACAGATGAGAGATGAGCA	4219
QY	1384	AATGATTAA 1392	
Db	4220	AATGATTAA 4228	
RESULT 5			
LOCUS	AR017561	1858 bp	DNA
DEFINITION	Sequence 6 from patent US 5780024.	linear	PAT 05-DEC-1998
ACCESSION	AR017561		
VERSION	AR017561.1	GI:3973164	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1858)		
TITLE	Brown, R.H., Fishman, P.S., Francis, J.W. and Hosler, B.A.		
JOURNAL	Superoxide dismutase/tetanus toxin fragment C hybrid protein		
FEATURES	Patent: US 5780024-A 6 14-JUL-1998;		
	Location/Qualifiers		
	1..1858		
	/organism="unknown"		
BASE COUNT	665 a	248 c	373 g
ORIGIN		572 t	
Query Match	99.1%;	Score 1379.8;	DB 6;
Best Local Similarity	99.9%;	Pred. No. 3.2e-173;	Length 1858;
Matches 1381;	Conservative 0;	Mismatches 2;	Indels 0;
			Gaps 0;
QY	10	TCACACCACAATTCACATTTCTATCTCTAAAAATCGAATTTGGGTGATATGAGAA	69
Db	476	TCACACCACACTCCATTTCTTATTCTTAAAAATCTGGAATTTGGGTGATATGAGAA	535
QY	70	GATATAGATGTTATATTTAAAAAAGATACAAATTTTAAATTTAGATATTAATATGATATT	129
Db	536	GATATAGATGTTATATTTAAAAAAGATACAAATTTTAAATTTAGATATTAATATGATATT	595
QY	130	ATATCAGATATATCTGGGTTTAATTCATCTGTAATATACATATCCAGATGCTCAATTGGTG	189
Db	596	ATATCAGATATATCTGGGTTTAATTCATCTGTAATATACATATCCAGATGCTCAATTGGTG	655
QY	190	CCCCGAATTAATGCAAGCAATACATTTAGTAAACAATGATCTTCTCAAGTTATAGTG	249
Db	656	CCCCGAATTAATGCAAGCAATACATTTAGTAAACAATGATCTTCTCAAGTTATAGTG	715

Qy	250	CATAAGCTATGAGATATTGCATATATACATATATGTTAAATAATTTACGGTTGACCTTTTGG	309
Db	716	CATAAAGCATGATGGATATTGCAATATATACATATATGTTAAATAATTTTACCGCTTACCTTTTGG	775
Qy	310	TTGAGGGTTCCCAAAAGATATCTGCTAGTCATTTAGAACATATGACCAATAGAGTATTTCA	369
Db	776	TTGAGGGTTCCCAAAAGTATCTGCTAGTCATTTAGAACATATGACCAATAGAGTATTTCA	835
Qy	370	ATAATTTAGCTCTATGAAAAAACATATGCTATCATATAGCATCTGTTTGGAGTGTATCACCT	429
Db	836	ATAATTTAGCTCTATGAAAAAACATATGCTATCATATAGCATCTGTTTGGAGTGTATCACCT	895
Qy	430	AAAGCTAATACTTATATTTGACCTTTAAAGAATTCGCGGGAGAACTTGAGACAAATTACT	489
Db	886	AAAGCTAATACTTATATATGACCTTTAAAGAATTCGCGGGAGAACTTGAGACAAATTACT	955
Qy	490	TTTAAAGGATTTTACCTGATTAATTTTAATGCTTATTTAGCAATAATAGGCTTTTATTAACCT	549
Db	956	TTTAAAGGATTTTACCTGATTAATTTTAATGCTTATTTAGCAATAATAGGCTTTTATTAACCT	1015
Qy	550	ATTACTAATGATAGATATTTATCTCTGCTAATTTGATATTAATAGGACTGTATTTAGGAAGCT	609
Db	1016	ATTACTAATGATAGATATTTATCTCTGCTAATTTGATATTAATAGGACTGTATTTAGGAAGCT	1075
Qy	610	GCAGAAATTTACTGTTTATAGACCTATTAGAGAGGATATATATTAACATTTAAACTAGAT	669
Db	1076	GCAGAAATTTACTGTTTATAGACCTATTAGAGAGGATATATATTAACATTTAAACTAGAT	1135
Qy	670	AGATGTAATATATATATATATCAATACGCTTCCTATTTGATATTAATAGATTTTGGCAAGA	729
Db	1136	AGATGTAATATATATATATATCAATACGCTTCCTATTTGATATTAATAGATTTTGGCAAGA	1195
Qy	730	TTTAAATCCAAAAGAGATTGAAAAATTTATACAGAGTTATTTATCTTAACCTTTTAAAG	789
Db	1196	TTTAAATCCAAAAGAGATTGAAAAATTTATACAGAGTTATTTATCTTAACCTTTTAAAG	1255
Qy	790	GACTTCTGGGGAACCCCTTTACGATATGATACAGATATTTATTTATACCAGTAGCTTCT	849
Db	1256	GACTTCTGGGGAACCCCTTTACGATATGATACAGATATTTATTTATACCAGTAGCTTCT	1315
Qy	850	ACGTCTAAAGATGTTTCATTTGAAAAATTAACAGATTAATATGATTTTGACAAATGGCCCA	909
Db	1316	ACGTCTAAAGATGTTTCATTTGAAAAATTAACAGATTAATATGATTTTGACAAATGGCCCA	1375
Qy	910	TGCTATCTACGAGAAATTTGAATATATATATATAGAGGTTATATTAATGAGCTAAAAATTT	969
Db	1376	TGCTATCTACGAGAAATTTGAATATATATATATAGAGGTTATATTAATGAGCTAAAAATTT	1435
Qy	970	ATTATTAAGAAGATATACCTCTATATATGAAATGATTCCTTTGTAAATAGGTGATTTT	1029
Db	1436	ATTATTAAGAAGATATACCTCTATATATGAAATGATTCCTTTGTAAATAGGTGATTTT	1495
Qy	1030	ATTAAATTTATGTTATCATATTAACAATTAATGAGCACATTTAGCTTATCCGAAAGATGGA	1089
Db	1496	ATTAAATTTATGTTATCATATTAACAATTAATGAGCACATTTAGCTTATCCGAAAGATGGA	1555
Qy	1090	AATGCTTAAATATATCTGTTGATTAATTTCTATAGAGTATAGGTTATATGCCCAGGATATCCCT	1149
Db	1556	AATGCTTAAATATATCTGTTGATTAATTTCTATAGAGTATATGATGCCCAGGATATCCCT	1615
Qy	1150	CTTTATTAATAAAAAATGAGACGATAAAAATTCGCTGATTTAAAAACCTATTTCTGTACAACTT	1209
Db	1616	CTTTATTAATAAAAAATGAGACGATAAAAATTCGCTGATTTAAAAACCTATTTCTGTACAACTT	1675
Qy	1210	AAATTTATATGATATATAAATATGATCTTTTGGAGCTAGTACCTACCCATATGAGCAATA	1269
Db	1676	AAATTTATATGATATATAAATATGATCTTTTGGAGCTAGTACCTACCCATATGAGCAATA	1735
Qy	1270	GGCAACGATCCAAATAGGAGTATATTTAAATTTGCAAGCAACATGGTACTTTTAATCAATTTAA	1329
Db	1736	GGCAACGATCCAAATAGGAGTATATTTAAATTTGCAAGCAACATGGTACTTTTAATCAATTTAA	1795
Qy	1330	GATTAATAATTTTAGAGTGTGATTTGGTACTTTGTACCTACATAGAGGATGGACAAATATGAT	1389

|||||
Db 1796 GATTAATTTAGATGTGTACTTTGTACCTACAGATGAAGATGACAAATGAT 1855
QY 1390 TAA 1392
|||
Db 1856 TAA 1858
RESULT 6
AF389424 3931 bp DNA linear BCT 26-JUL-2001
LOCUS Clostridium tetani tetanus toxin gene, partial cds.
DEFINITION AF389424
ACCESSION AF389424
VERSION AF389424.2 GI:15022164
KEYWORDS
SOURCE Clostridium tetani.
ORGANISM Clostridium tetani.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1 (bases 1 to 3931)
AUTHORS Shumin,Z. and Dianliang,L.
JOURNAL Cloning and sequence analysis of tetanus toxin gene
REFERENCE 2 (bases 1 to 3931)
AUTHORS Shumin,Z. and Dianliang,L.
JOURNAL Direct Submission
TITLE Submitted (04-JUN-2001) Department of Serum, National Institute for
the Control of Pharmaceutical and Biological Products, No. 2
Tiantanxili, Temple of Heaven, Beijing 100050, P.R. China
COMMENT On Jul 26, 2001 this sequence version replaced gi:14718605.
FEATURES
source location/Qualifiers
1..3931
/organism="Clostridium tetani"
/db_xref="taxon:1513"
<1..>3931
/codon_start=2
/transl_table=11
/product="tetanus toxin"
/protein_id="AAK72964.2"
/db_xref="GI:15022164"
/translation="RIPRINNRKSVNPVNDPIIMEPKCKDIIYKAKFTTORT
WVPERYEGTKPEDEPNPSSSLIEGASEYIDENYLTDSKDFLOTWKLFRIRKIN
VAGEMLDKITNAIPPLGNSYSLDEFDNNSVSNLEODSGPATVTSAMLTSLI
FGPGVLKNEVRGIVLRVDNKNYPFCRGFGSIMOMTCEPFDNVEMLTSLT
IGSKRYFDPDALLMHEILHVLHGLYGMVSSHEIIPSKOEIMOTVPIASAELEFE
GGODANLISIDIKNDLYEKTLDNYKAIAKLSOVTSCNPENIDISKOIYOXODP
KDSNGOXYVNEDEKPOLIVNSIMYGTEIEIGKPKNTKTLSSVSMHDPYKTNLDD
TLYNDEGPNIESKDLKSEYKGNKNRVNTNARNVDSGLVSKLICKKIIPPTNR
ENLIRKIALDLGELCLIKINEDLIFAEKNSFSEEPODEIVSNKPKPLNFY
SLDKIILLYNLOSKITLPPNDRTPYKRGIPYAPSEKSNASTIEIHINDONTIYOYLY
AOKSPITLORITMNSVDALINSTIYSPFVSVKNOGAGLFLQWRDIIIDE
TNESOKTIDKISDVSTIVPIYIPALNTIKOGYEGNFGALETTGVLLLEKIPET
LPVIALSIASSSTOKEKIKITIDNLEKREYEMIEVYLKVAKKWGLTVTQOKSY
QMRSLRYOVDAIKITIDYKITYSGPDEQLADETNINLNKLEEKANAMININIFM
RESSSPFLVNOMINETKOLEFDTQSKNIMQYIKANSFISITELKLESKINVE
STPIEPSFKNDICWDMEDIDILKSTIINLNDIIISIDFENSIVITYPDO
LVPNGKAIHLVNNSESEVYHKAMDINIDMNFENYSEFMRVKSASHLEPOYT
NEYSTISMKKYSLSIGSGVSIKGNNTLIMTLKDSAGVROTFPDLSPKRAYLAN
KWFETITNDRSSANLYINGVLMGSAETTGALREDNNTITKLDRCNNNOYVSD
KFRICKALNKEIEKLYTSYITFLRFWGPPLPYDEYVLIPIVYSSKDOQLNI
TDYMTLNAPSYTNGKRLNYRRLYSGLFIIKRYPNNEIDISFVSCGFIKLYSVN
NNEHIVGYPKDGNFNNDRILRVGNAPGIIPLKKEVAKGLDLKTYVSQULKYDOK
DASLGLVGHNGOIGNDPMDRDLILASNMVFNHLKDKTLCDWYFVPTD"

QY 64 GAAGAAGATATAGATGTATATTAATAAAGAGACATTTAAATTTAGATTTAAT 123
Db 2624 GAAGAAGATATAGATGTATATTAATAAAGAGACATTTAAATTTAGATTTAAT 263
QY 124 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAATAACATATCCAGATGCTCAA 183
Db 2684 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAATAACATATCCAGATGCTCAA 2743
QY 184 TTGGTCCCGGAAATTAATGCGAAGCAATATCAATTTGTAACATGAACTTCGAGAGTT 243
Db 2744 TTGGTCCCGGAAATTAATGCGAAGCAATATCAATTTGTAACATGAACTTCGAGAGTT 2803
QY 244 ATAGTGCATAAAGCTATGATATATGATATTAATGATATGTTTATATTTTACCGTTAGC 303
Db 2804 ATAGTGCATAAAGCTATGATATATGATATTAATGATATGTTTATATTTTACCGTTAGC 2863
QY 304 TTTTGTTGAGGGTTCCTTAAGTATCTGCTAGTCATTTAGACAAATATGACAAATGAG 363
Db 2864 TTTTGTTGAGGGTTCCTTAAGTATCTGCTAGTCATTTAGACAAATATGACAAATGAG 2923
QY 364 TATTCATATATATAGCTATATGAAAAACATAGCTATATAGATCTGGTGGAGTGA 423
Db 2924 TATTCATATATATAGCTATATGAAAAACATAGCTATATAGATCTGGTGGAGTGA 2983
QY 424 TCACCTTAAAGTAAATTAATTAATATGAGCTTTAAAGATCCCGGAGAGTTAGACAA 483
Db 2984 TCACCTTAAAGTAAATTAATTAATATGAGCTTTAAAGATCCCGGAGAGTTAGACAA 3043
QY 484 ATACTTTAGGAGTTTACCTGATATAATTAATGCTTATTTAGCAATAAATGAGGTTTT 543
Db 3044 ATACTTTAGGAGTTTACCTGATATAATTAATGCTTATTTAGCAATAAATGAGGTTTT 3103
QY 544 ATACTTATTAATATGATATGATTTATCTGCTAAATTTGATATTAATGAGTACTATG 603
Db 3104 ATACTTATTAATATGATATGATTTATCTGCTAAATTTGATATTAATGAGTACTATG 3163
QY 604 GGAATGCGAAGATTAATGATGTTTGGAGCTATTAAGAGAGATATATTAATTAATTA 663
Db 3164 GGAATGCGAAGATTAATGATGTTTGGAGCTATTAAGAGAGATATATTAATTAATTA 3223
QY 664 CTGATATGATGATTAATTAATTAATGATGATGATGATGATGATGATGATGATGAT 723
Db 3224 CTGATATGATGATTAATTAATTAATGATGATGATGATGATGATGATGATGATGAT 3283
QY 724 AAAGATTAATAATCCAAAGAGATTTGAAAAATTTATACCAAGTTATTTATACCTTT 783
Db 3284 AAAGATTAATAATCCAAAGAGATTTGAAAAATTTATACCAAGTTATTTATACCTTT 3343
QY 784 TTAAGAGCTTTGGGGAACCTTTAGCATATGATACGAAATATTTAATACCGTA 843
Db 3344 TTAAGAGCTTTGGGGAACCTTTAGCATATGATACGAAATATTTAATACCGTA 3403
QY 844 GCTTCTAGTCTTAAGATGTTCAATTAAGAAAAATTAACAGATTATGTTTACCAAT 903
Db 3404 GCTTCTAGTCTTAAGATGTTCAATTAAGAAAAATTAACAGATTATGTTTACCAAT 3463
QY 904 GCGCATGCTATACGAGAAATTTGAATATATATTAATGAAGTTATTAATGACACTA 963
Db 3464 GCGCATGCTATACGAGAAATTTGAATATATATTAATGAAGTTATTAATGACACTA 3523
QY 964 AAATTTATTAATAAAGATATACCTTAATGAATGATGATGATGATGATGATGATGAT 1023
Db 3524 AAATTTATTAATAAAGATATACCTTAATGAATGATGATGATGATGATGATGATGAT 3583
QY 1024 GATTTTATTAATTAATTAATGATATCAATTAATGAACATGATGATGATGATGATGATGAT 1083
Db 3584 GATTTTATTAATTAATTAATGATATCAATTAATGAACATGATGATGATGATGATGATGAT 3643
QY 1084 GATGAATATGCTTATTAATTAATCTTGAATGATGATGATGATGATGATGATGATGAT 1143
Db 3644 GATGAATATGCTTATTAATTAATCTTGAATGATGATGATGATGATGATGATGATGAT 3703

Oy	37	AAAAATCGGATTCGTGGGCTTGATAATGAAACAATATAGATGTATATATTAATAAACAAGT	96
Db	4	AAAATCTGGATCTGTGGGTTGATAATGAAGAACATATAGATGTATATTAATAAATAAGT	63
Oy	97	ACAAATTTAAATTTAGATTAATTAATGATATTTATACAGTATATCTGGGTTAATCA	156
Db	64	ACAAATTTAAATTTAGATTAATTAATGATATTAATATTCAGATATATCTGGGTTAATCA	123
Oy	157	TCTGTAATAACATATATCCAGATGCTCAATATGGTGGCCGGAATTAATGGCAAGCAATCAT	216
Db	124	TCTGTAATAACATATATCCAGATGCTCAATATGGTGGCCGGAATTAATGGCAAGCAATCAT	183
Oy	217	TTAGTAACAACATGATCTCTGGAAGTATATAGCATAAAGCTATGATATTTGAATATAT	276
Db	184	TTAGTAACAACATGATCTCTGGAAGTATATAGCATAAAGCTATGATATTTGAATATAT	243
Oy	277	GATATGTTAAATATTTTACCCTAGCTTTGGTTGAGGGTTCCTAAAGATCTGCTAGT	336
Db	244	GATATGTTAAATATTTTACCCTAGCTTTGGTTGAGGGTTCCTAAAGATCTGCTAGT	303
Oy	337	CATTTAGACAATATATGGCACAATAGATTAATCAATTAATATAGCTCTAGAAAAAACAATAGT	366
Db	304	CATTTAGACAATATATGGCACAATAGATTAATCAATTAATATAGCTCTAGAAAAAACAATAGT	363
Oy	337	CTATCAATAGATCTGGTTGGAGTGTATCACTTAAGGTAATTAATTAATATGACATTTA	456
Db	364	CTATCAATAGATCTGGTTGGAGTGTATCACTTAAGGTAATTAATTAATATGACATTTA	423
Oy	457	AAAGATTTCCGGGGAGAAGATTAGACAATTAACCTTTAGGCAATTACTGATTAATTTAAT	516

Db	424	AAAGATTCCGGGGAGAGAGTTAGACAAATACTTTAGGATTTACCGTAATAATTAT	483
QY	517	GCTTATTTAGCAATAAATGGGTTTTTATACATATTACTAATGATAGTTATCTTCTCT	576
Db	484	GCTTATTTAGCAAAATAAATGGGTTTTTATACATATTACTAATGATAGTTATCTTCTCT	543
QY	577	AATTTGTATTAATAATGGAGTACTTATGGGAAGTGCAGAAATTTACTGGTTTAGGACTATT	636
Db	544	AATTTGTATTAATAATGGAGTACTTATGGGAAGTGCAGAAATTTACTGGTTTAGGACTATT	603
QY	637	AGAGAGCATTAATTAATATACATTTAAACCTAGATAGTATTAATTAATAATCAATACGTT	696
Db	604	AGAGAGCATTAATTAATATACATTTAAACCTAGATAGTATTAATTAATAATCAATACGTT	663
QY	697	TCATTTGATTAATATTTAGGATATTTTTCGAAGCATTTAAATCCAAAAGACATTGAAAATTA	756
Db	664	TCATTTGATTAATATTTAGGATATTTTTCGAAGCATTTAAATCCAAAAGACATTGAAAATTA	723
QY	757	TACACAATTTTATTAATACATTAACCTTTTAAAGACCTCTGGGGAACCCCTTACGATAT	816
Db	724	TACACAATTTTATTAATACATTAACCTTTTAAAGACCTCTGGGGAACCCCTTACGATAT	783
QY	817	GATACAGATATTTATTTATATACAGTACCTCTAGTCTTAAAGATGTCAATTGAAAAAT	876
Db	784	GATACAGATATTTATTTATATACAGTACCTCTAGTCTTAAAGATGTCAATTGAAAAAT	843
QY	877	ATAACAGATTAATATGATTTTGACAAATGCCGCATCGTATCTAACGAAAAATTGATATTA	936
Db	844	ATAACAGATTAATATGATTTTGACAAATGCCGCATCGTATCTAACGAAAAATTGATATTA	903
QY	937	TATTTATTAAGGTTATATTAATGAGCTAAATTTTATTAATAAAGATATACACCTTAATAT	996
Db	904	TATTTATTAAGGTTATATTAATGAGCTAAATTTTATTAATAAAGATATACACCTTAATAT	963
QY	997	GAATATAGATTCCTTTGTTTAAATCAGCTGATTTTATTAATTAATATATGATATATACAT	1056
Db	964	GAATATAGATTCCTTTGTTTAAATCAGCTGATTTTATTAATTAATATATGATATATACAT	1023
QY	1057	AATGAGCATTGTAGTTATCCGAAGAATGGAAATGCCCTTAAATATCTTGATATGAATTT	1116
Db	1024	AATGAGCATTGTAGTTATCCGAAGAATGGAAATGCCCTTAAATATCTTGATATGAATTT	1083
QY	1117	CTAAGAGTAGGTTATTAATGCCCCAGGATCCCTCTTTATTAATAAATGGAAGGAGATATAA	1176
Db	1084	CTAAGAGTAGGTTATTAATGCCCCAGGATCCCTCTTTATTAATAAATGGAAGGAGATATAA	1143
QY	1177	TTGCGTGATTTAAATACCATCTCTATACACTTAAATTAATATGATGATTAATAATGCATCT	1236
Db	1144	TTGCGTGATTTAAATACCATCTCTATACACTTAAATTAATATGATGATTAATAATGCATCT	1203
QY	1237	TTAGGACTAGTAGTACCATTATGTCATTAATAGGCAACGATCCAAATAGGATATATTA	1296
Db	1204	TTAGGACTAGTAGTACCATTATGTCATTAATAGGCAACGATCCAAATAGGATATATTA	1263
QY	1297	ATTGCAAGCACTGCTACTTTTATATATTTAAAGTAAATTTTATGATGTGATTTGGTAC	1356
Db	1264	ATTGCAAGCACTGCTACTTTTATATATTTAAAGTAAATTTTATGATGTGATTTGGTAC	1323
QY	1357	TTTGTACCTACGATGAGGATGAGCAATATGATTA	1392
Db	1324	TTTGTACCTACGATGAGGATGAGCAATATGATTA	1359
RESULT 8			
128430			
LOCUS	128430	1359 bp	DNA
DEFINITION	Sequence 1 from patent US 5571694.		
ACCESSION	128430		linear
VERSION	128430.1	GI:1819206	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		

OLKNTDYMVLTNAPSTYNGKLNLYRRLYSGLKFLIKRYPNNEIDFVRSGDFIKL
YVSTRNNHHIYGPEDGNAPNNDBILPVGYNABPGLPKKEAVLRBLKRTYSVLK
LYDKDASLGLVGHNGOIGNDPNRDIILASNMVYRNLKDKLTICDMWIVPDEGWTN

BASE COUNT 525 a 147 c 225 g 459 t
ORIGIN

Query Match 96.0%; Score 1336.8; DB 1; Length 1356;
Best Local Similarity 99.1%; Pred. No. 1.7e-167;
Matches 1344; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

37 AAAAATCTGATTTGGTGTGATTAATGAAGAAGATAGATGTATATTAAAAAGAGT 96
1 AAAAATCTGATTTGGTGTGATTAATGAAGAAGATAGATGTATATTAAAAAGAGT 60
97 ACAATTTTAAATTTAGATATTAATAATGAATTAATATACAGATATATCTGGGTTTAATCA 156
61 ACAATTTTAAATTTAGATATTAATAATGAATTAATATACAGATATATCTGGGTTTAATCA 120
157 TCTGTAAATACATATCCATATGCTCAATGGTCCCGGAATTAATGGCAAGCATCAT 216
121 TCTGTAAATACATATCCATATGCTCAATGGTCCCGGAATTAATGGCAAGCATCAT 180
217 TTAGTAAACATGATCTCTGTAAGTTATAGTCATAAAGCTATGATTTGAATATAT 276
181 TTAGTAAACATGATCTCTGTAAGTTATAGTCATAAAGCTATGATTTGAATATAT 240
277 GATATGTTTAAATTTTACCCTAGCTTTGGTGGGTTCTTAAGATATCTGGTACT 336
241 GATATGTTTAAATTTTACCCTAGCTTTGGTGGGTTCTTAAGATATCTGGTACT 300
337 CATTTGACCAATATGTCACAAATGAGTATTTCAATAATTTAGCTATGAAAAACATAGT 396
301 CATTTGACCAATATGTCACAAATGAGTATTTCAATAATTTAGCTATGAAAAACATAGT 360
397 CATTCATAGATCTGTTGGAGTGTATCACTTAAGATTAATACTTAATATGAGCTTTA 456
361 CATTCATAGATCTGTTGGAGTGTATCACTTAAGATTAATACTTAATATGAGCTTTA 420
457 AAGATATCCGGGAGAGTATGACAAATACCTTTAGGGTTTACCTGATTAATTTAAT 516
421 AAGATATCCGGGAGAGTATGACAAATACCTTTAGGGTTTATCTGATTAATTTAAT 480
517 GCTTATTTAGCAATTAATGCGTTTATTAACATACTTAATGATAGATTAATCTCTGCT 576
481 GCTTATTTAGCAATTAATGCGTTTATTAACATACTTAATGATAGATTAATCTCTGCT 540
577 AATTTGATTAATTAATGAGTACTTATGGGAAGTCCAGAAATTTACTGTTAGGAGCTAT 636
541 AATTTGATTAATTAATGAGTACTTATGGGAAGTCCAGAAATTTACTGTTAGGAGCTAT 600
637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
601 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
697 TCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
661 TCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
757 TACACAAGTATTTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
721 TACACAAGTATTTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
781 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
877 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
841 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
937 TATATGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996

Db 901 TATATGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Qy 997 GAAATAGATTTCTTTGTTAAATCAGGTGATTTTATTAATTAATTAATTAATTAATTAAT 1056
Db 961 GAAATAGATTTCTTTGTTAAATCAGGTGATTTTATTAATTAATTAATTAATTAATTAAT 1020
Qy 1057 AATGAGCAGATTTAGTGTATTCGGAAGATGGAATGCGCTTTAATTAATCTTATAGATTT 1116
Db 1021 AATGAGCAGATTTAGTGTATTCGGAAGATGGAATGCGCTTTAATTAATCTTATAGATTT 1080
Qy 1117 CTAAGAGTGTATTAATGCCCCAGGTATCCCTTTTAAAAAAATGGAAGCAGATTAAT 1176
Db 1081 CTAAGAGTGTATTAATGCCCCAGGTATCCCTTTTAAAAAAATGGAAGCAGATTAAT 1140
Qy 1177 TTGGGTGATTTAAAAACCTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1236
Db 1141 TTGGGTGATTTAAAAACCTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1200
Qy 1237 TTAGGACTAGTGTATGACCAATATGCTCAATTAATTAATTAATTAATTAATTAATTAAT 1296
Db 1201 TTAGGACTAGTGTATGACCAATATGCTCAATTAATTAATTAATTAATTAATTAATTAAT 1260
Qy 1297 AATGCAACAGTGTGTATGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
Db 1261 AATGCAACAGTGTGTATGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
Qy 1357 TTGTGACCTAGATGAGGATGAGCAAAATGATTA 1392
Db 1321 TTGTGACCTAGATGAGGATGAGCAAAATGATTA 1356

RESULT 10
113981
LOCUS 113981 1359 bp DNA linear PAT 26-SEP-1995
DEFINITION Sequence 7 from patent US 5443966.
ACCESSION 113981
VERSION 113981.1 GI:996433
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1359)
AUTHORS Fairweather, N.F. and Makoff, A.J.
TITLE Expression of tetanus toxin fragment C
JOURNAL Patent: US 5443966-A 7 22-AUG-1995;
FEATURES
Location/Qualifiers
1..1359
/organism="unknown"

BASE COUNT 513 a 176 c 227 g 443 t
ORIGIN

Query Match 92.5%; Score 1287.2; DB 6; Length 1359;
Best Local Similarity 96.8%; Pred. No. 5.7e-161;
Matches 1313; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

37 AAAAATCTGATTTGGTGTGATTAATGAAGAAGATAGATGTATATTAAAAAGAGT 96
4 AAAAATCTGATTTGGTGTGATTAATGAAGAAGATAGATGTATATTCTGAAAAAGCT 63
97 ACAATTTTAAATTTAGATATTAATAATGAATTAATATACAGATATATCTGGGTTTAATCA 156
64 AACATTTCTGATTTGAGATCAACAGATATTAATCTCCGATCTCTGGTTCCAACCTC 123
157 TCTGTAAATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGGCAAGCAATCAT 216
124 TCTAGATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGGCAAGCAATCAT 183
217 TTAGTAAACATGATCTCTGTAAGTTATAGTCATAAAGCTATGATTAATTAATTAAT 276
184 TTAGTAAACATGATCTCTGTAAGTTATAGTCATAAAGCTATGATTAATTAATTAAT 243
277 GATATGTTTAAATTTTACCCTAGCTTTGGTGGGTTCTTAAGATATCTGGTACT 336

Db 244 GATATGTTTAAATTAACCGTTAGCTTTGGTTGAGGGTCCATAAGTATCTGCTAGT 303
 QY 337 CATTAGAACAAATATGGCAACAATGAGTATTCATATATAGTCTCTATGAAAAACATAGT 396
 Db 304 CATTAGAACAAATATGGCAACAATGAGTATTCATATATAGTCTCTATGAAAAACATAGT 363
 QY 397 CTAATCAATAGATTCGTTGGTGGAGTGTATCACTTAAGTAATAGCTTAATATGAGCTTTA 456
 Db 364 CTAATCAATAGATTCGTTGGTGGAGTGTATCACTTAAGTAATAGCTTAATATGAGCTTTA 423
 QY 457 AAAGATTCGCGGAGAACTTATAGCAAAATACCTTTAGGATTTAGCTGATTAATTTAAT 516
 Db 424 AAAGATTCGCGGAGAACTTATAGCAAAATACCTTTAGGATTTAGCTGATTAATTTAAT 483
 QY 517 GCTTATTTAGCAAAATTAATGAGTATTTTATACCTATTAATGATGATTAATGCTTCTGCT 576
 Db 484 GCTTATTTAGCAAAATTAATGAGTATTTTATACCTATTAATGATGATTAATGCTTCTGCT 543
 QY 577 AATTTGATTAATTAATGAGTATTTTATGGAAGTGGCAAAATTAATGAGTATTTAAT 636
 Db 544 AATTTGATTAATTAATGAGTATTTTATGGAAGTGGCAAAATTAATGAGTATTTAAT 603
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
 Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
 QY 697 TCTATTTAGTAAATTTAGGATTTTGGCAAGCATTAATCCAAAAGAGATTTGAAAATTA 756
 Db 664 TCTATTTAGTAAATTTAGGATTTTGGCAAGCATTAATCCAAAAGAGATTTGAAAATTA 723
 QY 757 TACCAAGTATTTATCTATTAATACCTTTTAAAGAGCTTGGGGAACCCCTTACGATAT 816
 Db 724 TACCAAGTATTTATCTATTAATACCTTTTAAAGAGCTTGGGGAACCCCTTACGATAT 783
 QY 817 GATACAGATTTATTTATTAATACAGTATCTATCTATTAAGATTTGATTAATTAATTAAT 876
 Db 784 GATACAGATTTATTTATTAATACAGTATCTATCTATTAAGATTTGATTAATTAATTAAT 843
 QY 877 ATACAGATTTATATGATTTATGACAAATGGCCAGTCGATTAACGGAATTTGATTAAT 936
 Db 844 ATACAGATTTATATGATTTATGACAAATGGCCAGTCGATTAACGGAATTTGATTAAT 903
 QY 937 TATATTTAGAGGTTATTAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 996
 Db 904 TATATTTAGAGGTTATTAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 963
 QY 997 GAAATAGATTTCTTTGTTAAATCAGTGATTTTATTAATTAATTAATTAATTAATTAAT 1056
 Db 964 GAAATAGATTTCTTTGTTAAATCAGTGATTTTATTAATTAATTAATTAATTAATTAAT 1023
 QY 1057 AATGAGCAATTTAGGTTATCCCAAAAGATGGAATGCTTTAATTAATGATTAATTAAT 1116
 Db 1024 AATGAGCAATTTAGGTTATCCCAAAAGATGGAATGCTTTAATTAATGATTAATTAATTAAT 1083
 QY 1117 CTAAGAGTATTAATTAATGAGTATCCCTCTTTATTAATTAATTAATTAATTAATTAAT 1176
 Db 1084 CTAAGAGTATTAATTAATGAGTATCCCTCTTTATTAATTAATTAATTAATTAATTAAT 1143
 QY 1177 TTGCGTATTTAAATTAATCTTCTGTAACAATTAATTAATTAATTAATTAATTAATTAAT 1236
 Db 1144 TTGCGTATTTAAATTAATCTTCTGTAACAATTAATTAATTAATTAATTAATTAATTAAT 1203
 QY 1237 TTGAGACTATAGTATCCCAAAATGATTAATGAGTATGAGTATTAATTAATTAATTAAT 1296
 Db 1204 TTGAGACTATAGTATCCCAAAATGATTAATGAGTATGAGTATTAATTAATTAATTAAT 1263
 QY 1297 ATTGCAAGCAACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
 Db 1264 ATTGCAAGCAACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
 QY 1357 TTTGTACCTACAGATGAGGATGAGCAAAATGATTAAT 1392
 Db 1324 TTTGTACCTACAGATGAGGATGAGCAAAATGATTAAT 1359

RESULT 11
 A18201 1359 bp DNA linear PART 09-JUL-2002
 LOCUS synthesised fragment C DNA increased (G+C) content seq ID No.2.
 DEFINITION
 ACCESSION A18201
 VERSION A18201.1 GI:212727091
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 DEFINITION
 artificial sequences.
 1 (bases 1 to 1359)
 AUTHORS
 Makoff,A.J., Romanos,M.A., Clare,J.J. and Fairweather,N.F.
 TITLE
 JOURNAL
 Patent: EP 0430645-A 3 05-JUN-1991;
 THE WELLCOME FOUNDATION LIMITED
 FEATURES
 source
 1..1359
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 381 a 374 c 271 g 333 t
 ORIGIN
 Query Match 57.4%; Score 799.2; DB 6; Length 1359;
 Best Local Similarity 74.3%; Pred. No. 1.2e-96;
 Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
 QY 37 AAAATCTGGATTGTGGGTTGATTAATGAGAAGATATAGATTAATTAATTAATTAATTAAT 96
 Db 4 AAAAACCTGATTTGTGGTGCACACGAAGAAGACATCGATTAATCTGAAAAAGTCT 63
 QY 97 ACAATTTAATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 156
 Db 64 ACCATCTGAACCTGGACATCAACACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 123
 QY 157 TCTGTATTAACATATCCAGATGCTCAATTTGGCCGGAATTAATTAATTAATTAATTAAT 216
 Db 124 TCTGTATTAACATATCCAGATGCTCAATTTGGCCGGAATTAATTAATTAATTAATTAAT 183
 QY 217 TTTGTAACATGATTAATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276
 Db 184 CTGGTTAACACAGATTTCTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 243
 QY 277 GATATCTTTAATTAATTTTACCGTTAGCTTTGTTGGAGGTTCCAAAGTATCTGCTAGT 336
 Db 244 GATATCTTTAATTAATTTTACCGTTAGCTTTGTTGGAGGTTCCAAAGTATCTGCTAGT 303
 QY 337 CATTAGAACAAATATGACCAAAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 396
 Db 304 CACTTGACAGTACGAGTACGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 363
 QY 397 CTAATCAATAGATTCGTTGGTGGAGTGTATCACTTAAGTAATAGCTTAATTAATGAGCTTTA 456
 Db 364 CTAATCAATAGATTCGTTGGTGGAGTGTATCACTTAAGTAATAGCTTAATTAATGAGCTTTA 423
 QY 457 AAAGATTCGCGGAGAACTTATAGCAAAATACCTTTAGGATTTAGCTGATTAATTTAAT 516
 Db 424 AAAGATTCGCGGAGAACTTATAGCAAAATACCTTTAGGATTTAGCTGATTAATTTAAT 483
 QY 517 GCTTATTTAGCAAAATTAATGAGTATTTTATACCTATTAATGATGATTAATGCTTCTGCT 576
 Db 484 GCTTATTTAGCAAAATTAATGAGTATTTTATACCTATTAATGATGATTAATGCTTCTGCT 543
 QY 577 AATTTGATTAATTAATGAGTATTTTATGGAAGTGGCAAAATTAATGAGTATTTAAT 636
 Db 544 AATTTGATTAATTAATGAGTATTTTATGGAAGTGGCAAAATTAATGAGTATTTAAT 603
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
 Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
 QY 697 TCTATTTGATTAATTTAGGATTTTGGCAAGCATTAATCCAAAAGAGATTTGAAAATTA 756

Db 1084 CTGCGTGTGGTTACACGCTCCGGGATCCCGCTGTACAAAAAATGAGACTGTTAA 1143

QY 1177 TTGGGTATTTAAACCTATCTGTACAACTTAATATATGATGATTAATGATCT 1236

Db 1144 CTGGGTGACGAAACCTACTCTGTACGCTGAAGTGTACGACGACAAACGCTTCT 1203

QY 1237 TTAGACTAGTAGTACCCTAATGATGATCAAAATAGCAAGATCCAAATAGGATATATTA 1296

Db 1204 CTGGGTGTGGTTGACCCACACGCTGATCGGTACGACCCGACCGCTGACATCTCG 1263

QY 1297 ATTGCAAGCACTGTACTTATCATTTAAAGTAAATTTTGGATGTGATGATC 1356

Db 1264 ATGCTTCTACTGTGTACTTACCACTGAAAGCAAAATCTGGGTGACCTGTAC 1323

QY 1357 TTGTACCTACAGATGAGGATGACAAATGATTAA 1392

Db 1324 TTGTTCCGACCGATGAAAGTTGAGCAACGACTTA 1359

RESULT 13

LOCUS A49987 3712 bp DNA circular PAT 07-MAR-1997

DEFINITION Sequence 4 from Patent WO9520655.

ACCESSION A49987

VERSION A49987.1 GI:2303172

KEYWORDS

SOURCE

ORGANISM

unidentified.

unclassified.

REFERENCE 1 (bases 1 to 3712)

AUTHORS Khan, M.A., Chatfield, S.N. and Li, J.

TITLE THE HTRA-PROMOTERS

JOURNAL EXPRESSION OF HETEROLOGOUS PROTEINS IN ATTENUATED BACTERIA USING

Medeva Holdings BY (NL)

Patient: WO 9520655-A 4 03-AUG-1995;

Other publication AU 1541595 950815.

COMMENT

FEATURES

location/Qualifiers

1..3712

BASE COUNT 972 a 967 c 841 g 932 t

ORIGIN

Query Match 57.4%; Score 799.2; DB 6; Length 3712;

Best Local Similarity 74.3%; Pred. No. 8.5e-97;

Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

Db 448 CTGTCCATCGGCTCTGGTTGCTGTTCCTGGAAGGTTAACAACTGATCGACTCG 507

QY 457 AAAGATCCCGGAGAGATAGACAAATATCTTTAGGATTTTCCGATTAATTTAT 516

Db 508 AAAGACTCCGGGCGAGAGTTCTGACATCTCTTCCGCGACTGCGGACAAAGTTCAAC 567

QY 517 GCTTATTTTACCAATTAATGGGTTTATTAACATTAATTAATGATAGATTAATCTCTCT 576

Db 568 GCGTACCTGGCTTAACAAATGGGTTTTCATCATCTAATGATGATCTCTCTCTCT 627

QY 577 AATTGTATTAATTAATGAGTACTATGAGGAAAGTGCAGAAATTAATGATGATGAT 636

Db 628 AACCTGTACATCAACGCGCTTCTGATGAGCTCGCTGTAATCACTGCTGCGCTATC 687

QY 637 AGAGGATTAATATATTAACATTAACATTAACATGATGATTAATTAATTAATCAATGCT 696

Db 688 CGTAGAGACAAACATCACTTAACTGATGATGATGATGATGATGATGATGATGAT 747

QY 697 TCTATGTATTAATTTAGATATTTTGAAGCATTAATCCAAAGATGGAATTA 756

Db 748 TCCATGACAAAGTTCCGATATCTTCTGCAAGACATGACCCGAAAGATGCAAAAACTG 807

QY 757 TACAGAGTTATTAATCTAATACCTTTTAAAGACTTGTGGGAAACCCCTTACGATAT 816

Db 808 TATACAGCTACCTCTCTATACCTTCTGCGGATCTTGGGTTAACCCGCTGCTTAC 867

QY 817 GATACAGATTAATTTATTAATTAACAGTACTTCTAGTTCTTAAGATGTTCAATGGAATAT 876

Db 868 GACACCGAATTAATACCTGATCCCGGATGCTTAAAGCTTAAAGACGTTGACGTGAAAAAC 927

QY 877 ATACAGATTTATTAATTTGACAAATGGGCAATGCTATCTTAACGAAATTAATTA 936

Db 928 ATCACTGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 987

QY 937 TATTATGAAGGTTATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 996

Db 988 TACTATCCAGCTCTGTACAAAGGCTGAAATTCATCATCAACGCTACACTCGGAACAC 1047

QY 997 GAATAGATTTCTTTTGTATTAATCAGTGATTTTATTAATTAATTAATTAATTAATTA 1056

Db 1048 GAATCGATTTCTTTTGTATTAATCAGTGATTTTATTAATTAATTAATTAATTAATTA 1107

QY 1057 AATGAGCACTTGAAGTTATCCGAAAGATGGAATGCTTTAATTAATTAATTAATTA 1116

Db 1108 AACGAAACATCTGTTGTTTACCCGAAAGCGTAAACGCTTTCAACAACTGACAGAAAT 1167

QY 1117 CTAAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1176

Db 1168 CTGGGTGTGGTTACAAAGCTCCCGGATATCCCGTGTACAAAAAATGAAAGCTGTTAA 1227

QY 1177 TTGGGTATTTAAACCTATCTGTACAACTTAATTAATTAATTAATTAATTAATTA 1236

Db 1228 CTGGGTATCTGAAACCTACTCTGTGATGATGATGATGATGATGATGATGATGATGAT 1287

QY 1237 TTAGACTAGTAGTACCCTAATGATGATCAAAATAGCAAGATCCAAATAGGATATATTA 1296

Db 1288 CTGGGTGTGGTTTACCAACAGCTGATGATGATGATGATGATGATGATGATGATGAT 1347

QY 1297 ATTGCAAGCACTGTACTTATCATTTAAAGTAAATTTTGGATGTGATGATC 1356

Db 1348 ATGCTTCTACTGTGTACTTACCACTGAAAGCAAAATCTGGGTGACCTGTAC 1407

QY 1357 TTGTACCTACAGATGAGGATGACAAATGATTAA 1392

Db 1408 TTGTTCCGACCGATGAAAGTTGAGCAACGACTTA 1443

RESULT 14

LOCUS A42484 4366 bp DNA linear PAT 06-MAR-1997

DEFINITION Sequence 12 from Patent WO9504151.

ACCESSION A42484

VERSION A42484.1 GI:2297933

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN

unidentified.
unclassified.
1 (bases 1 to 4366)
Khan, M.A., Hormeache, C.E., Chatfield, S.N. and Dougan, G.
VACCINE COMPOSITIONS EXPRESSION OF RECOMBINANT FUSION PROTEINS IN
ATTENUATED BACTERIA
Patent: WO 9504151-A 12 09-FEB-1995.
MEDEVA HOLDINGS B V (NL)
Other publication PL 313979 960805
Other publication GB 2295394 960529
Other publication JP 85036029 960423
Other publication CA 2168459 950209
Other publication NO 950348 950328
Other publication AU 7235794 950228
Other publication FI 950396 950130
Other publication AU 4719393 940303.
Location/Qualifiers
1..4366
/organism="unidentified"
/db.xref="taxon:32644"

Query Match 57.3%; Score 797.2; DB 6; Length 4366;
Best Local Similarity 74.3%; Pred. No. 1.5e-96;
Matches 1006; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

37 AAAAAATCGATTTGGTGGTGGATATATGACAGAAATATAGATGTTATATTAATAAAGAGT 96
101 AAAAACTTGTGTTGGTGGACAGCAAGAAAGATGATGTTATCTGTAAGAAAGTCT 160
97 ACAATTTAAATTTGATATTAATATATATATATATATATATATATATATATATATATAT 156
161 ACCATTTGAACTTGGACATCAACAGCAATATATATCTCGACATCTCGTTTCAACTCC 220
157 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGCAATATATGCAAAAGCAATACAT 216
221 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGCAATATATGCAAAAGCAATACAT 280
217 TTAGAAACATGATCTTCTGAAATTTAGTCAATTAAGCATGATGATGATGATGATGATGAT 276
281 CTGGTTAAACAGAACTTCTTGAATATATCTGCAACAAGCCATGACATGCAATACAAAC 340
277 GATATGTTAATATATTTACCGTTAGCTTTGGTGAAGGTTCTTAAAGTATCTGCTAGT 336
341 GACATGTTCAACAACTTCAACGTTAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 400
337 CATTTAAGCAATATGAGCAAAATGATTCATATATATATATATATATATATATATATATAT 396
401 CACCTGGAACATGAGCACTAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
397 CTATCAATAGATCTGTTGGAGTATCACTTAAAGTAATATATATATATATATATATATATAT 456
461 CTGTCATGCGGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 520
457 AAGATTTCCGCGGAGAAATAGACAATATATATATATATATATATATATATATATATATAT 516
521 AAGATTTCCGCGGAGAAATAGACAATATATATATATATATATATATATATATATATATAT 580
517 GGTATTTTAAAT 576
581 GGTATTTTAAAT 640
577 AATTTGAT 636
641 AACCTGAT 700
637 AAGAGGAT 696
701 CGTGAAGACAAACATCTCTTAAGCTGGAACGTTGCAACAAACAAACAAACAAACAAACAA 760

697 TCTATGTAAT 756
761 TCCATTCGCAAGATTCGCAATCTCTGCAAGCACTGAAACCCGAAAGATGAAATG 820
757 TACACAGATTTAT 816
821 TATACAGATTTAT 880
817 GATACAGATTTAT 876
881 GACACAGATTTAT 940
877 ATACAGATTTAT 936
941 ATACAGATTTAT 1000
937 TATATACAGATTTAT 996
1001 TACTACAGATTTAT 1060
997 GAATACAGATTTAT 1056
1061 GAATACAGATTTAT 1120
1057 AATACAGATTTAT 1116
1121 AATACAGATTTAT 1180
1117 CTAACAGATTTAT 1176
1181 CTGCTGATTTAT 1240
1177 TTGCTGATTTAT 1236
1241 CTGCTGATTTAT 1300
1237 TTAGAGATTTAT 1296
1301 CTGCTGATTTAT 1360
1297 AATGCAACATGTTAT 1356
1361 ATGCTGATTTAT 1420
1357 TTTGATTTAT 1390
1421 TTGCTGATTTAT 1454

RESULT 15
A37074
LOCUS A37074 3754 bp DNA circular PAT 05-MAR-1997
DEFINITION Sequence 17 from Patent WO9403615.
ACCESSION A37074
VERSION A37074.1 GI:2294248
KEYWORDS
ORGANISM
SOURCE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

unidentified.
unclassified.
1 (bases 1 to 3754)
Khan, M.A., Hormeache, C.E., Villarreal-Ramos, B., Chatfield, S.N. and
Dougan, G.
EXPRESSION OF RECOMBINANT FUSION PROTEINS IN ATTENUATED BACTERIA
Patent: WO 9403615-A 17 17-FEB-1994;
MEDEVA HOLDINGS B V (NL)
Other publication CA 2141477 940217
Other publication AU 719393 940303
Other publication NO 950348 950328
Other publication FI 950396 950130
Other publication JP 85036029 960423.
Location/Qualifiers
1..3754
/organism="unidentified"

BASE COUNT 978 a /db_xref="taxon:32644"
ORIGIN 977 c 860 g 939 t

Query Match 57.28; Score 796.8; DB 6; Length 3754;
Best Local Similarity 74.38; Pred. No. 1.8e-96;
Matches 1005; Conservative 0; Mismatches 347; Indels 0; Gaps 0;

QY 37 AAAAATCTGATTTGGTGTGATTAATGAGATATAGTGTATATTTAAAAAGAGT 96
Db 101 AAAAACCTGATTTGGTGTGACAGAGAGACATCGATTTATCCGAAAAAGTCT 160
QY 97 ACAATTTAAATTTAGATTAATTAATGATATTAATATGATATATGCGTTAAATTTCA 156
Db 161 ACAATTTGAACTTGGACATCAACAGATATTAATTCGCCGATCTGTGTTCACTCC 220
QY 157 TCTGTAATACATATCCAGATGCTCAATGTCGCCGATTAATGCAAGCAATATCAT 216
Db 221 TCTGTAATACATATCCAGATGCTCAATGTCGCCGATTAATGCAAGCAATATCAT 280
QY 217 TTGATTAACATATGATCTGTAAGTTATAGTCATTAAGCTATGATTTGATATAT 276
Db 281 CTGATTAACATATGATCTGTAAGTTATAGTCATTAAGCTATGATTTGATATAT 340
QY 277 GATATGTTAATTAATTTCCGTTAGCTTTGGTTGAGGTTCTTAAAGTATGCTAGT 336
Db 341 GACATGTTCAACACTTCCGTTAGCTTTGGTTGAGGTTCTTAAAGTATGCTAGT 400
QY 337 CATTTGAACAATATGACCAAAATGATATCAATTAATTAATTAATTAATTAAT 396
Db 401 CACCTGGAACAGTACGGCAGTAAAGCTACTCATCATGCTATGTAAGAACACTCC 460
QY 397 CTATCAATGAGATGCTGGTGGAGTATCACTTAAAGTAAATTAATTAATTAAT 456
Db 461 CTGTCATGCGCTCTGTTGTTGTTCCGTAAGGTTAAACACTGATGAGTCTG 520
QY 457 AAAGATTCGGGGAGAGTTAGACAAATTAATTTAGGATTTAOCGTATTAATTAAT 516
Db 521 AAAGATTCGGGGAGAGTTAGACAAATTAATTTAGGATTTAOCGTATTAATTAAT 580
QY 517 GCTTATTTAGCAAAATTAATGCTTTTAACTAATTAATTAATTAATTAATTAAT 576
Db 581 GCTTATTTAGCAAAATTAATGCTTTTAACTAATTAATTAATTAATTAATTAAT 640
QY 577 AATTTGATTAATTAATGAGTACTTATGGAAGTGCAGAAATTAAGTCTTAAAGTAT 636
Db 641 AACCTGATCAACAGGGCTTCTGATGGCTCCGCTGAATCACTGCTGGCGCTATC 700
QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
Db 701 CGTGAAGCAACATCACTTAAGCTGAGCGTTGCAACAAACACACAGTACGTA 760
QY 697 TCTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
Db 761 TCCATGCAAGTCCCTATCTTCTGCAAAAGCACTGAACCGGAAGATCGAAAACTG 820
QY 757 TACCAAGTATTTATCTATTAACCTTTTAAAGACTTCTGGGAAACCCCTTACGATAT 816
Db 821 TATACAGCTACCTGCTATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
Db 881 GACACCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 940
QY 877 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
Db 941 ATCAGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
QY 937 TATTTAGAGGTTATATATATGACTTAAATTAATTAATTAATTAATTAATTAAT 996
Db 1001 TACTACCGAGCTGTGTACAGCGCTGAATTAATTAATTAATTAATTAATTAAT 1060
QY 997 GAAATAGATTTCTTTGTTAAATCAGGTGATTTTATTAATTAATTAATTAATTAAT 1056

Db 1061 GAAATCGATTTCTTTGTTAAATCTGTAATCACTGATCACTCACTGATCTTACAAAC 1120
QY 1057 AATGAGCAATTTGATTTATCCGAAAGATGGAATGCTTAAATTAATTAATTAAT 1116
Db 1121 AACGAAACATCTGTTGTTACCGGAAAGACGTTACAGCTTCAACACTGACAGAA 1180
QY 1117 TTAAGAGTATTTAAATGCCCCGAGTATCCCTCTTTTAAAAAATGAAAGCAATAA 1176
Db 1181 CTGGGTGTTGTTTACACGCTCCGGGTATCCGCTGTACAAAAAATGAAAGCTGTAAA 1240
QY 1177 TTGCTGATTTAAACCTATTTCTGACAACTTAATTAATTAATTAATTAATTAAT 1236
Db 1241 CTGGGTGATCGAAGAACTTCTGTTGAGTGAACCTGTAACGACCCGAAACCTGATCCT 1300
QY 1237 TTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
Db 1301 CTGGGTGTTGTTGTTACCAACGAGTACGATCGTAAACGACCCGAAACCTGATCCT 1360
QY 1297 ATTGCAAGCACTGTTACTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
Db 1361 ATGCTTCTTAACGTGTTCTTCAACCACTGTAAGACAAATTAATTAATTAATTAAT 1420
QY 1357 TTGTTACTACATGAGGATGACAAATGA 1388
Db 1421 TTGTTCTGACCGATGAAGGTTGACCAACGA 1452

Search completed: December 1, 2002, 09:49:39
Job time : 2749 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:53:27 ; Search time 254 Seconds

(without alignments)
12341.666 Million cell updates/sec

Title: US-09-816-467-3

Sequence: 1 atggttttttcaacccaat.....aggatggaacaatgtaa 1392

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392	100.0	1600	20 AAX27234	Tetanus toxin frag
2	1387.4	99.7	1851	8 AAN70545	Sequence encoding
3	1379.8	99.1	1859	19 AAV2580	DNA encoding a SOD
4	799.2	57.4	1359	12 AA012121	Synthetic tetanus
5	797.2	57.4	3712	16 AAO97490	Plasmid pHTRA1. N
6	796.8	57.3	4366	16 AA085424	Plasmid pTECH3-p28
7	796.8	57.2	3754	15 AAO57880	Intermediate plasm
8	796.8	57.2	3754	16 AAO5420	Plasmid pTECH1. S
9	796.8	57.2	3769	15 AAO57881	Intermediate plasm

10	796.8	57.2	3769	16 AAO97492	Plasmid pTECH2. N
11	796.8	57.2	4377	16 AAO85421	Plasmid pTECH1-p28
12	521	37.4	1766	16 AAO97491	Plasmid pTECH2 + h
13	248.2	17.8	1535	19 AAV30596	Clostridium botuli
14	234.8	16.9	1547	19 AAV30597	Clostridium botuli
15	233.2	16.8	3876	19 AAV30599	Clostridium botuli
16	232.8	16.7	1526	19 AAV30580	Clostridium botuli
17	228.2	16.4	1317	21 AA287220	DNA encoding nativ
18	228.2	16.4	1546	19 AAV30575	Clostridium botuli
19	228.2	16.4	2532	17 AA287218	DNA encoding nativ
20	228.2	16.4	3891	17 AA287218	DNA encoding nativ
21	228.2	16.4	4835	21 AA287218	DNA encoding nativ
22	220.8	15.9	4835	21 AA287218	DNA encoding nativ
23	210.2	15.1	1472	19 AAV26289	Clostridium botuli
24	206.2	14.8	1293	18 AAV30585	Clostridium botuli
25	205.2	14.7	4017	22 AAF58862	Immunogenic type F
26	205	14.7	1463	19 AAV30584	Clostridium botuli
27	176.2	12.7	1460	19 AAV30593	Clostridium botuli
28	163.6	11.8	1917	21 AA287219	DNA encoding nativ
29	152.6	11.0	1469	19 AAV30591	DNA encoding nativ
30	113.8	8.2	3950	20 AAX23521	Clostridium botuli
31	109	7.8	1502	19 AAV30588	DNA encoding nativ
32	93.8	6.7	1278	22 AAV30588	DNA encoding nativ
33	93.8	6.7	1400	22 AAV30588	DNA encoding nativ
34	93.8	6.7	1400	22 AAV30588	DNA encoding nativ
35	85	6.1	1830	24 ABL56243	Human immune syste
36	85	6.1	3292	24 ABL56203	Human immune syste
37	84.2	6.0	15548	24 ABL34155	Human immune syste
38	75.6	5.4	1511	13 AAO28302	Human immune syste
39	75.6	5.4	1511	15 AAO66798	Human immune syste
40	75.6	5.4	1511	19 AAV14508	Human immune syste
41	75.6	5.4	1511	20 AAV10082	Human immune syste
42	75.6	5.4	15548	24 ABL34155	Human immune syste
43	75.6	5.4	5000	24 ABL35643	Human immune syste
44	74.8	5.4	6175	24 ABL35307	Human immune syste
45	74.6	5.4	50000	24 ABL56201	Human immune syste

ALIGNMENTS

RESULT 1	AAX27234 standard; DNA; 1600 BP.
ID	AAX27234
AC	AAX27234;
XX	
XX	
DF	28-MAY-1999 (first entry)
XX	
DE	Tetanus toxin fragment C coding sequence.
KW	Tetanus toxin fragment C; TTC; central nervous system; CNS; spinal cord; proteolytic fragment; retrograde axonal transport; spinal cord disease; transsynaptic transport; neurodegenerative disease; motoneuron disease; amyotrophic lateral sclerosis; spinal muscular atrophy; therapy; ALS; SMA; neurodegenerative lysosomal storage disease; neuronal maping; ss.
XX	
OS	Clostridium tetani.
XX	
PN	MO9909057-A2.
XX	
PD	25-FEB-1999.
XX	
PF	12-AUG-1998; 98MO-EP05113.
XX	
PR	13-NOV-1997; 97US-0065236.
XX	
PR	14-AUG-1997; 97US-0055615.
XX	
PA	(INSP) INST PASTEUR.
XX	
PI	Brulet P, Coen L, Osta Pinzolas R,
XX	
DR	WPI; 1999-180971/15.

DR P-PSDB: AAY00921.

XX Delivery of a composition to the central nervous system or spinal
cord - comprises administration of a non-toxic, proteolytic
fragment of tetanus toxin in association with a molecule having
biological function

PS Example 1; Fig 1; 53bp; English.

XX This sequence encodes the tetanus toxin fragment C (TTC).
CC The invention relates to a method for in vivo delivery of a desired
CC composition into a human or animal central nervous system (CNS) or spinal
CC cord comprising administering a non-toxic, proteolytic fragment of tetanus
CC toxin (TTC) in association with at least a molecule having a biological
CC function and where the composition is capable of in vivo retrograde
CC axonal transport and transsynaptic transport into the CNS or the spinal
CC cord of the human or animal and of being delivered to different areas of
CC the CNS or the spinal cord. The method can be used for the treatment of
CC humans or animals with CNS or spinal cord disease, e.g. neurodegenerative
CC and motoneuron diseases such as amyotrophic lateral sclerosis (ALS),
CC spinal muscular atrophies (SMA) or neurodegenerative lysosomal storage
CC diseases. Compositions comprising hybrid fragments of TTC comprising
CC fragments C and B can also be used for neuronal mapping and
CC immunisations. Use of TTC comprising fragments A, B and C results in
CC better transport of the fragment inside the organism compared with
CC fragment C.

XX Sequence 1600 BP; 589 A; 213 C; 278 G; 520 T; 0 other;

Query Match 100.0%; Score 1392; DB 20; Length 1600;
Best Local Similarity 100.0%; Pred. No. 4.3e-230;

Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTTTTCACACCAATTCATTTCTTATTTCTAATAAATTCGATGTTGGGTTGAT 60
DB ATGCTTTTTCACACCAATTCATTTCTTATTTCTAATAAATTCGATGTTGGGTTGAT 147
QY 61 AATGAGAGAGATATGATGATATTTAAAGAGTACATTTTAAATTTGATATTAAT 120
DB AATGAGAGAGATATGATGATATTTAAAGAGTACATTTTAAATTTGATATTAAT 207
QY 148 AATGAGAGAGATATGATGATATTTAAAGAGTACATTTTAAATTTGATATTAAT 207
DB AATGAGAGAGATATGATGATATTTAAAGAGTACATTTTAAATTTGATATTAAT 207
QY 121 AATGATATTTATGATGATATTTCTGGTTTAAATTCATCTGTAATATTCAGATCT 180
DB AATGATATTTATGATGATATTTCTGGTTTAAATTCATCTGTAATATTCAGATCT 267
QY 208 AATGATATTTATGATGATATTTCTGGTTTAAATTCATCTGTAATATTCAGATCT 267
DB AATGATATTTATGATGATATTTCTGGTTTAAATTCATCTGTAATATTCAGATCT 267
QY 181 CAATTGGTGGCCGGAATTAATGGCAACCAATACATTTAGTAACAATGAAATCTTGAA 240
DB CAATTGGTGGCCGGAATTAATGGCAACCAATACATTTAGTAACAATGAAATCTTGAA 327
QY 241 GTTATAGTGCATTAAGCTATGATATTTGAATTAATGATATTTAAATTTTACCGTT 300
DB GTTATAGTGCATTAAGCTATGATATTTGAATTAATGATATTTAAATTTTACCGTT 387
QY 328 GTTATAGTGCATTAAGCTATGATATTTGAATTAATGATATTTAAATTTTACCGTT 387
DB GTTATAGTGCATTAAGCTATGATATTTGAATTAATGATATTTAAATTTTACCGTT 387
QY 301 ACCTTTTGGTTGAGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 360
DB ACCTTTTGGTTGAGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 447
QY 388 ACCTTTTGGTTGAGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 447
DB ACCTTTTGGTTGAGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 447
QY 361 GAGTATTTCAATTAATGCTATGATGATGATGATGATGATGATGATGATGATGAT 420
DB GAGTATTTCAATTAATGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 448 GAGTATTTCAATTAATGCTATGATGATGATGATGATGATGATGATGATGATGAT 507
DB GAGTATTTCAATTAATGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 421 GATACACTTAAGGTAATTAATTAATGATGATGATGATGATGATGATGATGATGAT 480
DB GATACACTTAAGGTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 508 GATACACTTAAGGTAATTAATTAATGATGATGATGATGATGATGATGATGATGAT 567
DB GATACACTTAAGGTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 481 CAATTAATCTTTAGGATTTACCTGATTAATTAATTAATGATGATGATGATGATGAT 540
DB CAATTAATCTTTAGGATTTACCTGATTAATTAATTAATGATGATGATGATGATGATGAT 627
QY 568 CAATTAATCTTTAGGATTTACCTGATTAATTAATTAATGATGATGATGATGATGAT 627
DB CAATTAATCTTTAGGATTTACCTGATTAATTAATTAATGATGATGATGATGATGATGAT 627
QY 541 TTTTATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB TTTTATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
QY 628 TTTTATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
DB TTTTATTAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687

QY 601 ATGGAGAGTGCAGAAATTAATCTGTTAGGAGCTATTAGAGAGATTAATATTAACATTA 650
DB ATGGAGAGTGCAGAAATTAATCTGTTAGGAGCTATTAGAGAGATTAATATTAACATTA 747
QY 661 AAACATGATAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB AAACATGATAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 807
QY 721 TGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB TGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 867
QY 781 TTTTATTAAGAGCTTTGGGAGAACCTTTTACGATGATGATGATGATGATGATGATGAT 840
DB TTTTATTAAGAGCTTTGGGAGAACCTTTTACGATGATGATGATGATGATGATGATGAT 927
QY 841 GTAGCTTTGATGTTCTAAGATGTTCAATTTGAAAAATTAACAGATTTATGATTTGACA 900
DB GTAGCTTTGATGTTCTAAGATGTTCAATTTGAAAAATTAACAGATTTATGATTTGACA 987
QY 901 AATGGCCGATCGTATTAACGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATGA 960
DB AATGGCCGATCGTATTAACGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATGA 1047
QY 961 CTAAATTTTATTAATAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
DB CTAAATTTTATTAATAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1107
QY 1021 GGTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
DB GGTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1167
QY 1108 GGTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
DB GGTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1227
QY 1168 AAAGATGAAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
DB AAAGATGAAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1287
QY 1228 GGTATCCCTCTTTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
DB GGTATCCCTCTTTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1347
QY 1201 GTACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB GTACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1407
QY 1261 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1467
QY 1321 CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1300
DB CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1387
QY 1408 CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1467
QY 1381 ACAATGATTA 1392
DB ACAATGATTA 1479
QY 1468 ACAATGATTA 1479
DB ACAATGATTA 1479

RESULT 2
AAN70545
ID AAN70545 standard; DNA; 1851 BP.
AAN70545;
AC AAN70545;
XX
XX 22-APR-1991 (first entry)
XX
DE Sequence encoding a portion of the B fragment and all of the C
DE fragment of tetanus toxin.
XX
XX TT; vaccine; ds.
XX
XX Clostridium tetani.
XX
XX
FH Key Location/Qualifiers

```

FT      CDS      1..1722
XX      /*lag= a
XX      EP209281-A.
XX      PD      21-JAN-1987.
XX      PF      27-JUN-1986; 86EP-0305029.
XX      PR      28-JUN-1985; 85GB-0016442.
XX      (WELL ) WELLCOME FOUNDATION LTD.
XX      PI      Fairweather NF;
XX      DR      WPI; 1987-015999/03.
XX      P-PSDB; AAP70345.
XX      PT      Cloned DNA sequence coding for tetanus toxin - or its fragments
XX      contg. epitope used to express antigens for vaccine production.
XX      PS      Claim 4; Fig 1; 36pp; English.
XX      CC      Gene product comprises a tetanus toxin fragment, which may be
XX      CC      expressed in a transformed host, and used as an antigen in vaccine
XX      CC      production, against the disease.
XX      SO      Sequence 1851 BP; 754 A; 204 C; 279 G; 614 T; 0 other;

Query Match      99.7%; Score 1387.4; DB 8; Length 1851;
Best Local Similarity 99.9%; Pred. No. 2.7e-229;
Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

YY      4      GTTTTTCACACCAATTCATTTCTTATTCCTAAATAATCTGATGTTGGTGGATAT 63
DB      334      GTTTTTCACACCAATTCATTTCTTATTCCTAAATAATCTGATGTTGGTGGATAT 393
YY      64      GAAGAATATGATGTTATATTAATAAGATACATTTTAAATTTAGATTTATAT 123
DB      394      GAAGAATATGATGTTATATTAATAAGATACATTTTAAATTTAGATTTATAT 453
YY      124      GATATTTATATCAGATATATCTGGTTTAAATTCCTGTAATTAACATATCCAGATCTCAA 183
DB      454      GATATTTATCAGATATATCTGGTTTAAATTCCTGTAATTAACATATCCAGATCTCAA 513
YY      184      TTGGTCCCGGGAATTAATGCAAGCAATTCATTTAGTAACAATGATCTTGGAAGTT 243
DB      514      TTGGTCCCGGGAATTAATGCAAGCAATTCATTTAGTAACAATGATCTTGGAAGTT 573
YY      244      ATATGTCATTAAGCTATGATATTTGAATATGATATGTTTATATATTTTACCGTTAGC 303
DB      574      ATATGTCATTAAGCTATGATATTTGAATATGATATGTTTATATATTTTACCGTTAGC 633
YY      304      TTTTGGTTAGAGGTTCTTAAGATATCTGCTAGCATTTAGAACAAATATGACAAATGAG 363
DB      634      TTTTGGTTAGAGGTTCTTAAGATATCTGCTAGCATTTAGAACAAATATGACAAATGAG 693
YY      364      TATTCATTAATTAAGCTCTTGAATAAACAATAGCTATCAATAGATCTGTTGGAGTGA 423
DB      694      TATTCATTAATTAAGCTCTTGAATAAACAATAGCTATCAATAGATCTGTTGGAGTGA 753
YY      424      TCACCTTAAGGTAATTAATTAATGAGCTTTAAAGATTCGCGGGAGAGATTAGCAAA 483
DB      754      TCACCTTAAGGTAATTAATTAATGAGCTTTAAAGATTCGCGGGAGAGATTAGCAAA 813
YY      484      ATAACTTTTAGGATTTACCTGATTAATTAATGCTTATTTAGCAATAAATGGGTTTTT 543
DB      814      ATAACTTTTAGGATTTACCTGATTAATTAATGCTTATTTAGCAATAAATGGGTTTTT 873
YY      544      ATAACTTTTAGGATTTACCTGATTTCTGCTAATTTTATATTAATGAGATCTTATG 603
DB      874      ATAACTTTTAGGATTTACCTGATTTCTGCTAATTTTATATTAATGAGATCTTATG 933

```

```

YY      604      GGAAGTCAGAAATTTACTGGTTAGAGCTATTAGAGGATTAATATTAACATTTAAA 663
DB      934      GGAAGTCAGAAATTTACTGGTTAGAGCTATTAGAGGATTAATATTAACATTTAAA 993
YY      664      CTAGATGATGTAATTAATTAATCAATAGCTTCTATTTAGTAATTTAGATTTTTC 723
DB      994      CTAGATGATGTAATTAATTAATCAATAGCTTCTATTTAGTAATTTAGATTTTTC 1053
YY      724      AAGCATTTAATCCAAAGAGATTGAAAATTTATACACAGTTATTTATCATTAACCTT 783
DB      1054      AAGCATTTAATCCAAAGAGATTGAAAATTTATACACAGTTATTTATCATTAACCTT 1113
YY      784      TTAGAGACTTCTGGGAAACCCCTTACGATATGATGATATTAATTTATTAATCCAGTA 843
DB      1114      TTAGAGACTTCTGGGAAACCCCTTACGATATGATGATATTAATTTATTAATCCAGTA 1173
YY      844      GCTTCTAGTTCTTAAGATGTTCAATTTGAAAATTAACAGATTAATTAATTTGACAAAT 903
DB      1174      GCTTCTAGTTCTTAAGATGTTCAATTTGAAAATTAACAGATTAATTAATTTGACAAAT 1233
YY      904      GCGCCATCGTATCTAACGGAATTTGAATTAATATTAATTAATTAATTAATTAATTA 963
DB      1234      GCGCCATCGTATCTAACGGAATTTGAATTAATATTAATTAATTAATTAATTAATTA 1293
YY      964      AATTTTATTAATAAAGATATACACCTTAATTAATTAATTAATTTGTTAAATCAGGT 1023
DB      1294      AATTTTATTAATAAAGATATATACACCTTAATTAATTAATTAATTTGTTAAATCAGGT 1353
YY      1024      GATTTTATTAATTAATTAATGATATATTAACATTAATTAATTAATTAATTAATTAAT 1083
DB      1354      GATTTTATTAATTAATTAATGATATATTAACATTAATTAATTAATTAATTAATTAAT 1413
YY      1084      GATGGAATGCTTTATATATCTTGAAGATTTCTAAGATTAATTAATTAATTAATTAAT 1143
DB      1414      GATGGAATGCTTTATATATCTTGAAGATTTCTAAGATTAATTAATTAATTAATTAAT 1473
YY      1144      ATCCCTCTTTATTAATAAATAATGGAAGCAGTAATAATTCGATTTTAAATACCTATTCGTA 1203
DB      1474      ATCCCTCTTTATTAATAAATAATGGAAGCAGTAATAATTCGATTTTAAATACCTATTCGTA 1533
YY      1204      CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
DB      1534      CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593
YY      1264      CAATATGCAACGATCAATATAGGATATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
DB      1594      CAATATGCAACGATCAATATAGGATATTAATTAATTAATTAATTAATTAATTAATTAAT 1653
YY      1324      TTAATAAGATTAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
DB      1654      TTAATAAGATTAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1713
YY      1384      AATGATTA 1392
DB      1714      AATGATTA 1722

RESULT 3
AAV32580
ID      AAV32580 standard; DNA; 1858 BP.
XX
XX      AAV32580;
XX
XX      23-SEP-1998 (first entry)
XX
XX      DNA encoding a SOD-1/TTC hybrid protein.
XX
XX      Chimeric; copper-zinc superoxide dismutase; SOD-1; TTC; SOD.1et451;
XX      tetanus toxin fragment C; tetanus holotoxin; nerve cell; stroke;
XX      neurological disorder; oxidative stress; brain hypoxia-reperfusion;
XX      epilepsy; Parkinson's disease; Huntington's disease; ss.
XX
OS      Chimeric - Homo sapiens.

```

```

OS Chimeric - Clostridium tetani.
XX Key Location/Qualifiers
FH CDS 2..1858
FT /tag= a
FT /product= "SOD-1/TTC hybrid protein"
FT 2..490
FT /tag= b
FT /note= "encodes SOD-1 region of the SOD-1/TTC
FT hybrid protein"
FT misc_feature 503..1855
FT /tag= c
FT /note= "encodes TTC region of the SOD-1/TTC
FT hybrid protein"
XX US5780024-A.
XX 14-JUL-1998.
XX 21-JUN-1996; 96US-0668381.
XX 23-JUN-1995; 95US-0000473.
XX 21-JUN-1996; 96US-0668381.
XX (GENE ) GEN HOSPITAL CORP.
XX (ORMA-) UNIV MARYLAND BALTIMORE.
XX Brown RH, Fishman PS, Francis JW, Hosler BA:
XX WPI: 1998-412999/35.
XX P-PSDB; AAM48909.
XX
XX New hybrid protein of superoxide dismutase and tetanus toxin
XX fragment C - having increased uptake by neurons and retention of
XX enzymatic activity in these cells, for treating neurological
XX diseases associated with oxidative stress
XX
XX Disclousure; Columns 25-28; 23pp; English.
XX
XX The present DNA sequence encodes an enzymatically active human
XX copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl
XX terminus with the tetanus toxin fragment C (TTC) moiety. The TTC
XX moiety constitutes amino acid residues 865-1315 of the tetanus
XX holotoxin. The hybrid protein, referred as SOD:tet451, is claimed
XX to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic
XX activity; (b) the TTC moiety selectively binds to nerve cells and
XX allows uptake of the hybrid protein into these cells; and (c) it
XX retains substantial SOD enzymatic activity following cellular uptake.
XX SOD:tet451 is claimed to be useful for treating neurological disorders
XX associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion,
XX epilepsy, Parkinson's and Huntington's diseases.
XX
XX Sequence 1858 BP; 665 A; 248 C; 373 G; 572 T; 0 other:
SQ
Query Match 99.1%; Score 1379.8; DB 19; Length 1858;
Best Local Similarity 99.9%; Pred. No. 5.3e-228;
Matches 1381; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 250 CATAAAGCTATGATATGTAATATATATGATATGTTAATATTTTACCGTTAGCTTTGG 309
DB 716 CATAAAGCTATGATATGTAATATATGATATGTTAATATTTTACCGTTAGCTTTGG 775
QY 310 TTGAGGTTCCCTAAAGATATCTGCTAGTCATTTAGAACAAATATGGCAAAATGAGTATCA 369
DB 776 TTGAGGTTCCCTAAAGATATCTGCTAGTCATTTAGAACAAATATGAGTATGAGTATCA 835
QY 370 ATATATGCTCTATGAAAAAACAATAGCTATATCATATGAGATCTGGTGGAGTATCATCTT 429
DB 836 ATATATGCTCTATGAAAAAACAATAGCTATATCATATGAGATCTGGTGGAGTATCATCTT 895
QY 430 AAGGTAAATACCTTAATATGAGACTTTAAAGATTCGCGGAGAAAGTTAGACAAATTAAT 489
DB 896 AAGGTAAATACCTTAATATGAGACTTTAAAGATTCGCGGAGAAAGTTAGACAAATTAAT 955
QY 490 TTTAGGATTTACCTGATAAATTTAATGCTTATTTAGCAAAATAAATGGTTTTATTAAT 549
DB 956 TTTAGGATTTACCTGATAAATTTAATGCTTATTTAGCAAAATAAATGGTTTTATTAAT 1015
QY 550 ATTACTAATGATATGATTAATCTTCTGCTAATTTGTATATAATGAGTATGAGTATGAGT 609
DB 1016 ATTACTAATGATATGATTAATCTTCTGCTAATTTGTATATAATGAGTATGAGTATGAGT 1075
QY 610 GCAGAAATTAAGCTTTAGAGAGCTATTTAGAGAGATATATATTAACATTAAACCTAGAT 669
DB 1076 GCAGAAATTAAGCTTTAGAGAGCTATTTAGAGAGATATATATTAACATTAAACCTAGAT 1135
QY 670 AGATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT 729
DB 1136 AGATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT 1195
QY 730 TTAATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789
DB 1196 TTAATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
QY 790 GACTTCGGGGAACCCCTTACGATATGATATGATATGATATATATATATATATATATAT 849
DB 1256 GACTTCGGGGAACCCCTTACGATATGATATGATATATATATATATATATATATATATAT 1315
QY 850 AGTTCTTAAGATTTTCAATTTGAAAAATATATACAGATATATATATATATATATATATAT 909
DB 1316 AGTTCTTAAGATTTTCAATTTGAAAAATATATACAGATATATATATATATATATATATAT 1375
QY 910 TCGTATCTAACGGAATATATATATATATATATATATATATATATATATATATATATAT 969
DB 1376 TCGTATCTAACGGAATATATATATATATATATATATATATATATATATATATATATAT 1435
QY 970 ATTATATAAAGATATACACCTTAATATGAAATAGATTCTTTGTTAAATCAGGTGATTTT 1029
DB 1436 ATTATATAAAGATATACACCTTAATATGAAATAGATTCTTTGTTAAATCAGGTGATTTT 1495
QY 1030 ATTATATAAAGATATATATATATATATATATATATATATATATATATATATATATAT 1089
DB 1496 ATTATATAAAGATATATATATATATATATATATATATATATATATATATATATATAT 1555
QY 1090 AATGCCCTTAATATCTTGTATAGAAATTTCTAAGAGTAGTTAATGCCAGATATCCCT 1149
DB 1556 AATGCCCTTAATATCTTGTATAGAAATTTCTAAGAGTAGTTAATGCCAGATATCCCT 1615
QY 1150 CTTTATATAAATATGAAAGAGTAAATTTGGGTGATTTAAAAACCTATCTGTACAACTT 1209
DB 1616 CTTTATATAAATATGAAAGAGTAAATTTGGGTGATTTAAAAACCTATCTGTACAACTT 1675
QY 1210 AAATATATGATGATTAATAAATATGATCTTTAGAGTAGTATACCATTAATGGTCAATA 1269
DB 1676 AAATATATGATGATTAATAAATATGATCTTTAGAGTAGTATACCATTAATGGTCAATA 1735
QY 1270 GGCAGCATCCAAATAGAGATATATTAATTTGCAAGCAACTGTACTTTAATCATTTTAA 1329
DB 1736 GGCAGCATCCAAATAGAGATATATTAATTTGCAAGCAACTGTACTTTAATCATTTTAA 1795

```



```

Db      1204 CTGGGCTGCTGGTGGTACCCACCAACGGGTGAGTGGTAACGCCGCGTGCATCCG 1263
QY      1297 ATTGCAGCACTGGTACTTATATCATTTAAAGATAAATTAGATGTGATGTGAC 1356
Db      1264 ATCGCTTCTACTAGCTGACTTCAACACCGTGAAGACAAATTCCTGGTTGCCACTGGTAC 1323
QY      1357 TTTGTACTACAGATGAGGATGACAAATGATTAA 1392
Db      1324 TTCGTTCCGACGATGAGTGGTGGACCAAGACTAA 1359

RESULT 5
AA097490
ID      AA097490 standard; DNA; 3712 BP.
XX
AC      AA097490;
XX
DT      08-DEC-1995 (first entry)
XX
DE      Plasmid pHTRA1.
XX
KM      htra promoter; vaccine; attenuation; vector; tetanus toxin;
KW      pHTRA1; ds.
XX
OS      Not specified.
XX
FH      Key
FT      promoter
FT      Location/Qualifiers
FT      1..55
FT      /tag= a
FT      /label= htra_promoter
XX
PN      W09520665-A.
XX
PD      03-AUG-1995.
XX
PF      31-JAN-1995; 95WO-GB00196.
XX
PR      31-JAN-1994; 94GB-0001795.
XX
PA      (MEDE-) MEDEVA HOLDINGS BV.
XX
PI      Chatfield SN, Khan MA, Li J;
XX
WP: 1995-275450/36.
XX
DR      New DNA constructs contg. the htra promoter - used partic. for inducing
XX      heterologous protein prodn. in vivo, esp. in vaccine compositions
XX
PS      Disclosure; Page 28-31; pp: English.
XX
CC      Vector plasmid pHTRA1 contains the S. typhimurium temp.-inducible
CC      htra promoter sequence linked to the gene coding for tetanus toxin
CC      fragment C (tet-C). The vector is used to express Tet-C in
CC      attenuated Salmonella strains for vaccine prodn.
XX
SQ      Sequence 3712 BP; 972 A; 967 C; 841 G; 932 T; 0 other;

Query Match 57.4%; Score 799.2; DB 16; Length 3712;
Best Local Similarity 74.3%; Pred. No. 1.9e-128;
Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
QY      37 AAAAAATCGATTTGGTGGTATGATGAAGATATATGATATATTAATAAAGAGT 96
Db      88 AAAAAATCGATTTGGTGGTATGATGAAGATATATGATATATTAATAAAGAGT 147
QY      97 ACAAATTTAAATTTAGATATTAATATGATATATATATATATCTGGGTTTAATTCA 156
Db      148 ACCATTTCTGACCTGGACATCAACAACGATATTATCTCCGACATCTCTGGTTCAACTCC 207
QY      157 TCTGTATATACATATTCGACATGCTCAATTTGGTGGGATTAAGTCAAGCAATATCAT 216
Db      208 TCTGTATATCAATATTCGACATGCTCAATTTGGTGGGATTAAGTCAAGCAATATCAT 267

```

```

QY      217 TTAGTAACATGATCTTCTGAGTTATAGTCATTAAGCTATGATATTAATAT 276
Db      268 CTGGTTAAACAGAAATCTTCTGAAGTTATGTGACAAAGGCCATGACATCATACAAAC 327
QY      277 GATATGTTAAATTAATTTTACCGTTTGGTCTTTGGTGAAGGTTCTTAAGTATCTGCTAGT 336
Db      328 GACATTTTCAACAACCTTCAACCGTTTACTCTTCTGGCTGGCGGTTCCGAAGTTCTGCTTCC 387
QY      337 CATTTAGAACATATATGCGCAAAATGAGTATTCATTAATTAATAGCTCTATGAAGAAACATAGT 396
Db      388 CACCTGGAAACAGTACGCGACATAAGGATCTCATCATGCTCTATGAAGAAACATCTCC 447
QY      397 CTATCAATAGGATCTGGTGGATGATACCTTAAGGATTAATTAATATGAGACTTTA 456
Db      448 CTGTCCATCTGGCTGTGGTTGGTGTCTGTTTCCCTGAAGGTTAACAACCGATCTGGACTCTG 507
QY      457 AAGATTTCCGCGGAGAGTGTAGCAAAATACCTTTAGGATTTTACTGTATTAATTTAAT 516
Db      508 AAGACTCCGCGGCGCAAGTTGTCAGATCATTCCGCGACCTGCGGACAAAGTTCAAC 567
QY      517 GCTTATTTACCAATTAATGGGTTTTTATACATTAATTAATGATATATCTTCTGCT 576
Db      568 GCGTACCTGGCTAAACAATGGGTTTTTCATCATCTACATACATGCTGTCTTCTGCT 627
QY      577 AATTTGATATTAATGAGATCTTATGGGAAGTGAGAAATTAATCTGTTTGAAGACTAAT 636
Db      628 AACTGTACATCAAGCGCGTTCTGATGGGTCGCGTAATCAGTGGTGGCGCTATAC 687
QY      637 AGAGAGATTAATTAATTAATCACTTAATAACCTAGATATGATATTAATTAATCAATACGTT 696
Db      688 CGTGGGACAACAACATCATCTTACCTTACCTGAGCGGTGCAACAACAACACAGTACTGTA 747
QY      697 TCTATGATTAATTTAGGATATTTTGCAGCAAGCTTAATCCCAAGAGATTTGAAAAATTA 756
Db      748 TCCATGCAAGATTTCCGATCTTCTGCAAGCACTGAACCGGAAAGATGAAAAACTG 807
QY      757 TACCAAGTTATTTATCTATACCTTTTAAAGACTTTCTGGGAAACCTTTTACGATAT 816
Db      808 TATACGACCTACCTGCTATACCTTCCGCGTACCTTCCGCGTACCTGCGGTAC 867
QY      817 GATACAGAAATTTATTTATTAATCCAGTACTCTGTTTAAAGATGTTCAATTTGAAAAAT 876
Db      868 GACACCGAATATTTACCTGATCCCGGTACTCTTACTTAAGAGCTCAGCTGAAAAAC 927
QY      877 ATACAGATTTATGATATTTGACAACATGCGCATCTGATACACGAAATTAATTAAT 936
Db      928 ATCACTGACTACATGTACTTACCAACGCGCTCTACACTAACGTAACGTGAACATC 987
QY      937 TATTATAGAGTTATATTAATGAGCTAAATTTATTAATAAAGATATACCTAATAT 996
Db      988 TACTTACCGACGCTGTACACGCGCTGAATTCATCATCAACGCTACACCTCGAAACAC 1047
QY      997 GAAATGATTTCTTTGTTAAATCAGTGATTTTATTAATTAATTAATGATATTAATTAAT 1056
Db      1048 GAAATGATTTCTTTGTTAAATCAGTGATTTTATTAATTAATTAATGATATTAATTAAT 1107
QY      1057 AATGAGCAATTTAGGTTATTCGAAAGATGGAATGCTTTAATTAATCTGATTAAT 1116
Db      1108 AAGCAACACATGCTGTTTACCGGAAAGCGGTATACCGCTTTAACAACAACCTGGAAGAAAT 1167
QY      1117 CTAAAGATGATTAATTAATGCGCGAGTATCCCTTTAATAAATAAATGAGAACAGTAATA 1176
Db      1168 CTGCGTGTGGTTTACAAACCTCCGGGTATCCCGCTGTACAAAAAATGGAACGTTTAA 1227
QY      1177 TTGCGGATTTAAAAACCTATTCTGTACAACTTAATTAATTAATGATATTAATTAAT 1236
Db      1228 CTGCGGATCTGAAAAACCTATCTGTTACGTAAGAACTTAAGACGCAAAAAACGTTCT 1287
QY      1237 TTAGCACTGTAGTATACCATTAATGCTCAATTAAGCAAGATCAATTAAGGATATATTA 1296
Db      1288 CTGGTCTGTGGTGTGACCAACGCTGATGCTGATGAGACCGCAACCTGACATCTG 1347

```


QY 1297 ATTGACAGCAACGTCGTAATCTTAAATCAATTTAAAGATAAATTTAGATGTGATTGGTAC 1356
 Db 1348 ATGCGCTTCTAAGTGTGATCTTCAACACCTGAAAGACAAATCTGGGTTGGGACTGGTAC 1407
 QY 1357 TTGTACCTACAGATGAGGATGAGCAATGATTA 1392
 Db 1408 TTGCTTCCGACCGATGAGGTTGACCAAGACACTAA 1443

RESULT 6
 AA085424
 ID AA085424 standard; DNA; 4366 BP.
 AC AA085424;
 XX
 DT 04-SEP-1995 (first entry)
 DE Plasmid pTECH3-P28.
 XX
 KW M13B; promoter; tetanus C toxin; TetC; fusion protein; vaccine;
 KW Plasmid pTECH3-P28; P28; glutathione-S-transferase;
 KW Schistosoma mansoni; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 98..2095
 FT /*tag= a
 PN WO9504151-A.
 XX
 PD 09-FEB-1995.
 PF 29-JUL-1994; 94WO-GB01647.
 XX
 PR 30-JUL-1993; 93WO-GB01617.
 PR 31-JAN-1994; 94GB-0001787.
 XX
 PA (MEDE-) MEDEVA HOLDINGS BV.
 PI Chatfield SN, Dougan G, Hormaeche CE, Khan MA;
 DR WPI: 1995-082234/11.
 XX
 PT DNA encoding fusion protein of tetanus toxin C fragment and
 PT heterologous antigen - does not contain connecting hinge region,
 PT useful in human and animal vaccines.
 PS Disclosure; Page 32-35; 44pp; English.
 CC Plasmid pTECH3-P28 (AA085424) was constructed that provided
 CC stable expression of a soluble fusion protein comprising
 CC the S. mansoni P28 protein as a C-terminal fusion to fragment
 CC C of tetanus toxin, without a connecting hinge region. The
 CC fusion protein was expressed in *Salmonella typhimurium* SL5338
 CC and recovered by glutathione-agarose affinity purification.
 SQ Sequence 4366 BP; 1176 A; 1081 C; 1011 G; 1098 T; 0 other;

Query Match 57.38; Score 797.2; DB 16; Length 4366;
 Best Local Similarity 74.38; Pred. No. 4.2e-128;
 Matches 1006; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

QY 37 AAAAATCTGATTTGGTGTGATTAATGAGAGATATAGATGTTATTAATAAAAGAGT 96
 Db 101 AAAAATCTGATTTGGTGTGATTAATGAGAGATATAGATGTTATTAATAAAAGAGT 96
 QY 97 ACAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 156
 Db 161 ACAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 156
 QY 157 TCTGTATAACATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 216
 Db 157 TCTGTATAACATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 216

Db 221 TCTGTATACATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 280
 QY 217 TTAGTAACATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 276
 Db 281 CTGCTTAACAAAGATCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 340
 QY 277 GATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 336
 Db 341 GATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 340
 QY 337 CAATTAAGCAATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 396
 Db 401 CAATTAAGCAATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 460
 QY 397 CTATCAATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 456
 Db 461 CTATCAATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 520
 QY 457 AAAGATTTCCGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
 Db 521 AAAGATTTCCGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
 QY 517 GCTTATTTAGCAATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 576
 Db 581 GCTTATTTAGCAATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 640
 QY 577 AATTGCTATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 636
 Db 641 AATTGCTATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 700
 QY 637 AGAGAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 696
 Db 701 AGAGAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 760
 QY 697 TCTATTTGATTAATTAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 756
 Db 761 TCTATTTGATTAATTAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 820
 QY 757 TACACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 816
 Db 821 TACACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 880
 QY 817 GATACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 876
 Db 881 GATACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 940
 QY 877 ATAACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 936
 Db 941 ATAACAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1000
 QY 937 TATATAGAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 996
 Db 1001 TATATAGAGATATATATATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1060
 QY 997 GAAATGATTTTGTGTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1056
 Db 1061 GAAATGATTTTGTGTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1120
 QY 1057 AATGAGACATTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1116
 Db 1121 AATGAGACATTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1180
 QY 1117 CTAGAGATGATTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1176
 Db 1181 CTAGAGATGATTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1240
 QY 1177 TTGGTGATTTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1236
 Db 1241 TTGGTGATTTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1300
 QY 1237 TTAGAGATGATTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1296
 Db 1301 TTAGAGATGATTTAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1360

QY 1237 TTAGACTAGTAGTACCCATATGTCATATAGGCAACGATCCAAATAGGAGATATATTA 1296
 Db 1301 CTGGGTCTGGTGTGATCCACACAGGTCAGATCGTAAACGACGACCCGACATCCTG 1360
 QY 1297 ATTGCAAGCACTGTACTTATATCATTTAAAGATTAATTTAGATGTGATTTGATC 1336
 Db 1361 ATGCTCTTAACTGTCTACTTCAACCACTGAAAGACAAATCTGCGTGTGACGTGAC 1420
 QY 1357 TTGTACTTACAGATGAGGATGACCAATGA 1388
 Db 1421 TTGCTCCGACCGATGAAGTTGGACCAACGA 1452

RESULT 8

AA085420
 ID AA085420 standard; DNA: 3754 BP.

AC AA085420;

DT 04-SEP-1995 (first entry)

DE Plasmid pTECH1.

KW NirB, promoter; tetanus C toxin; fusion protein; vaccine;
 plasmid pTECH1; ss.

OS Synthetic.

PN WO9504151-A.

PD 09-FEB-1995.

PF 29-JUL-1994; 94WO-CB01647.

PR 30-JUL-1993; 93WO-CB01617.

PR 31-JAN-1994; 94GB-0001787.

PA (MEDE-) MEDEVA HOLDINGS BV.

PI Chaffield SN, Dougan G, Hormaeche CE, Khan MA;
 WPI: 1995-082234/11.

PT DNA encoding fusion protein of tetanus toxin C fragment and
 heterologous antigen - does not contain connecting hinge region,
 useful in human and animal vaccines.

PS Disclosure; Page 25-27; 44pp; English.

CC Plasmid pTECH1r15, encoding tetanus toxin C from a nirB promoter,
 was used to construct the pTECH1 plasmid (given in AA085420) that
 incorporated a polylinker region suitable as a site for insertion
 of heterologous DNA to direct the expression of fragment C fusion
 proteins, e.g. the Schistosoma mansoni P28 antigen.

CC Sequence 3754 BP: 978 A; 977 C; 860 G; 939 T; 0 other;

Query Match 57.2%; Score 796.8; DB 16; Length 3754;
 Best Local Similarity 74.3%; Pred. No. 4, 9e-128;
 Matches 1005; Conservative 0; Mismatches 347; Indels 0; Gaps 0;

QY 37 AAAAATCTGATTTCTGGTTGATATGAAGAAGATATAGATGTTATATTAATAAAGAT 96
 Db 101 AAAAATCTGATTTCTGGTTGATATGAAGAAGATATAGATGTTATATTAATAAAGAT 160
 QY 97 ACAATTTAATTTAGATATTAATTAATGATATATATCAATATATCTGGTTAATTTCA 156
 Db 161 ACAATTTAATTTAGATATTAATTAATGATATATATCAATATATCTGGTTAATTTCA 220
 QY 157 TCTGATTAATCAATATCAATATGCTCAATTTGTCGCGAATTAATGCAAGAATCAAT 216
 Db 221 TCTGATTAATCAATATCAATATGCTCAATTTGTCGCGAATTAATGCAAGAATCAAT 280

QY 217 TTAGTAACAATGAATCTTCTGAAAGTATAGTCATTAAGATAGATGATATGATATAT 276
 Db 281 CTGGTTAACAAGATCTTCTGAAAGTATAGTCATTAAGATAGATGATATGATATAT 340
 QY 277 GATATGTTAATTAATTTAGCGTTAGCTTTTGGTTGAGGTTCTTAAGATATCTGATG 336
 Db 341 GACATGTTCAACAACTTCAACCGTTAGCTTTGAGGTTCTTAAGATATCTGATG 400
 QY 337 CATTTAGAACAAATATGCAACAATGAGTATTCATTAATTTAGCTATGAAACATATG 396
 Db 401 CACCTGGAACAGTACGGCACTAAGAGATACCTCATCACTCTATGAAACATATG 460
 QY 397 CTATCAATAGATGCTGTTGAGAGTATCACTTAAGGTAATTAATTAATGAGCTTGA 456
 Db 461 CTGTCATCGGCTCTGTTGAGTGTCTGTTCCCTGAAGGTTAACAACATGATGACCTG 520
 QY 457 AAAGATTCGCGGAGAGTAGACAATATCAATTTTGGGTTTACCTGATTAATTTAT 516
 Db 521 AAAGATTCGCGGAGAGTAGACAATATCAATTTTGGGTTTACCTGATTAATTTAT 580
 QY 517 GCTTATTTAGCAATTAATGAGGTTTATTAATTAATTAATTAATTAATTAATTTAT 576
 Db 581 GCGTACCTGGCTTAACAATATGAGTGTTCATCACTATCACTAAGATGCTGCTGCT 640
 QY 577 AATTGTATTAATTAATGAGTATTAATTAATTAATTAATTAATTAATTAATTTAT 636
 Db 641 AACCTGATCAATCAACGCGCTTCTGATGAGGCTCGCTGAATTAATCACTGCTGCT 700
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 696
 Db 701 CGTGAAGCAACAACATCACTTTAAGCTGAGCGTTGCAACAACAACGATGCTGA 760
 QY 697 TCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 756
 Db 761 TCAATGACAAGTTCGTTATCTTCTGCAAAAGCACTGAACCGCAAGATGCAAAAC 820
 QY 757 TACACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 816
 Db 821 TATACAGCTACCTGCTATCACTTCCCTGCTGAGGTTAGGCTGCTGCTGCT 880
 QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 876
 Db 881 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 940
 QY 877 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 936
 Db 941 ATCACTGATCACTATCACTGATCACTGATCACTGATCACTGATCACTGATCA 1000
 QY 937 TATTAAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 996
 Db 1001 TACTACCGACGCTGTACACAGGCTGAATTCATCACTACCTGCAACAC 1060
 QY 997 GAAATAGATTTCTTTGTTAATCAAGTATTTATTAATTAATTAATTAATTTAT 1056
 Db 1061 GAAATAGATTTCTTTGTTAATCAAGTATTTATTAATTAATTAATTAATTTAT 1120
 QY 1057 AATGAGCATTTGATTTATTCGCAAGATGAAATGCTTTAATTAATTTATTTAT 1116
 Db 1121 AATGAGCATTTGATTTATTCGCAAGATGAAATGCTTTAATTAATTTATTTAT 1180
 QY 1117 CTAAGATAGGTTATTAATGCCCCAGATATCCCTTTAATTAATTAATTAATTTAT 1176
 Db 1181 CTAAGATAGGTTATTAATGCCCCAGATATCCCTTTAATTAATTAATTAATTTAT 1240
 QY 1177 TTGGGTATTTAATAAACCTTTCTGTACAACTTAATTAATTTATTTATTTATTTAT 1236
 Db 1241 CTGGGTATTTAATAAACCTTTCTGTACAACTTAATTAATTTATTTATTTATTTAT 1300
 QY 1237 TTAGGACTAGTAGTACCAATTAATGTCATAATAGCAAGTCCAAATTAAGGATATTTA 1296
 Db 1301 CTGGGTCTGTTGTTAGTACCAACAGGTCATAGTGTAAACGCGAAGCTGATCTG 1360
 QY 1297 ATTGCAAGCAACTGTACTTATATCATTTAAAGATTAATTTAGATGTGATTTGATC 1356

QY	1237	TTTGAGCTACTAGTACGCCCAATGTCGAATTAGCGAACGATCCAAATAGGATATATTA	1296
Db	1301	CTGGGTCTGGTGTGGTACCACAACGGTCAGATCGGTAAAGCAGCCGAGCATCTCG	1360
QY	1297	ATTGCAGAAGCACTGACTTAAATCATTTAAAGATATAATTTAGAGATGATGGTGAC	1356
Db	1361	ATCGCTTCTAAGTGGTACTTCCAACCCCTGAAGACAAAAATCCTGGGTGGCGACTGTAC	1420
QY	1357	TTCGTACTTACATGACAGTAGGAGTGAGCAATATGA	1388
Db	1421	TTCGTTCCGACCGATGAAAGTTGAGCCAAACGA	1452
 RESULT 10 AAQ97492			
ID	AAQ97492	standard; DNA; 3769 BP.	
XX	AAQ97492;		
AC			
XX			
DT	08-DEC-1995	(first entry)	
XX			
DE	Plasmid pTECH2.		
KW	htxA promoter; vaccine; attenuation; vector; tetanus toxin;		
KW	pHTRAZ; pHTRCH2; ds.		
OS	Not specified.		
XX			
PN	W09520665-A.		
PD	03-AUG-1995.		
XX			
PF	31-JAN-1995;	95WO-GB00196.	
XX			
PR	31-JAN-1994;	94GB-0001795.	
XX			
PA	(MEDEV) MEDEVIA HOLDINGS BV.		
XX			
PI	Chatfield SN, Khan MA, Li J;		
DR	WPI; 1995-275450/36.		
XX			
PT	New DNA constructs contg. the htxA promoter - used partic. for inducing		
XX	heterologous protein prodn. in vivo, esp. in vaccine compositions		
XX			
PS	Disclosure: Page 31-34; pp: English.		
CC	Vector plasmid pHTRAL contains the S. typhimurium temp.-inducible		
CC	htxA promoter sequence linked to the gene coding for tetanus toxin		
CC	fragment C (het-C). It was used to prepare pHTRAZ in which a		
CC	hinge region is present at the C-terminal of the Tet-C fragment.		
CC	This hinge region is obld. from pTECH2.		
XX			
SQ	Sequence 3769 BP; 982 A; 980 C; 863 G; 944 T; 0 other:		
 Query Match 57.2%; Score 796.8; DB 16; Length 3769; Best Local Similarity 74.3%; Pred. No. 4.9e-128; Matches 1005; Conservative 0; Mismatches 347; Indels 0; Gaps 0;			
QY	37	AAAATCTGAGTTGGTGGTGGATATGAAAGATATAGATGTATATTAANAAGAGT	96
Db	101	AAAAMCCTTGATGTGGGTGCGACACAGCAAGACATGATGTATCTCGAAAAAGTCT	160
QY	97	ACAATTTTAAATTTAGATATTTAATAATGATATTATATCAGATATATCTGGCTTAATTC	156
Db	161	ACCAATTTGACCTTGGACATCAACACAGATATTATCTCGACATCTCGGTTTCACTCC	220
QY	157	TCTGTAATAACATNTCCAGATGCTCAATTTGGTCCCGCAATTAATGGCAAAACAATACAT	216
Db	221	TCTGTATACATATCCAGATGCTCAATTTGGTCCCGCATCAACGCAAAAGCTATCCAC	280
QY	217	TTAGTAACAATGATGATCTTCTGGAAGTTATATGATGATTAAGGTATGAGATATTGAATATAT	276

Db	281	CTGGTTAAACAAGAAATCTTCTGAAGTTATCGTGCAACAAAGGCATGGACATTCGATTCACAC	340
Qy	277	GATATGTTTATATATTTTACCGTTAGCTTTTGGTTGAAGGTTCTPAAAGTATCGTCTAGT	336
Db	341	GACATGTTTCAACAACCTTCACCGTTAGCTTCGGCGCGCGCTTCCAAAGTTCTCGTTCC	400
Qy	337	CATTTAACAATAATGTCACAAATGAGATTCATTAATTAGCTCTATGAAAAAACACTACT	396
Db	401	CACCTGGAACAGTAGCGGACACTAGAGAGACTCCATCAGCTCTATGAAAGAACCTCC	460
Qy	397	CTATCAATAGAGATCGTTGGAGTGTACTTAAAGSTAATTACTTAATATGACTTTA	456
Db	461	CTGTCCATCGGCTCGTTGGTCTGTTCCTCGTAAGGGTAAACACTATCTGCACTCTG	520
Qy	457	AAAGATTTCCCGGAGAGATTAGCAAAATACTTTTAGGATTTACCGTATTAATTTAT	516
Db	521	AAAGACTCGCGCGGAGAGTTCCGCAGATCTTTCCGCGACTCGCGGACAAAGTTTCAAC	580
Qy	517	GCTTATTTAGCAATAAATGGGTTTTATACTATTACTAATGATAGATTATCTTGTCT	576
Db	581	GGCTACTGCGCTAAACAATGGTTTTATCTATCACTATCAAGATCGTGTCTGTCTGT	640
Qy	577	AATTGTATTAATGAGTACTATGAGGAAGTGCAGAAATTACTGGTTTAGAGACTATT	636
Db	641	AACCTGTACTCAACGGCGTTCTGATGGCTCCGCTGAAATCAGTGGTCTGGCGCTATC	700
Qy	637	AGAGAGATTAATAATATACATTAAAACGTAGATGTATTAATTAATATCAATACGTT	696
Db	701	CGTAGAGACAACAACATCACTCTPAACTGAGCGTTGCAACAACAACACAGTACGTA	760
Qy	697	TCATTTGATTAATTTAGGATATTTTGCAAAAGCATTAATCCAAAAGAGATTGAAAAATTA	756
Db	761	TCCATTCGACAAGTTCGATACCTTCTGCAAAAGCACTGAACCGGAAAGAGATCGAAAACTG	820
Qy	757	TACACAAGTTATTTATCTATACCTTTTAAAGACTCTGCGGGAACCCCTTGAGAT	816
Db	821	TTACACAGTACCTGTACTACCTTCCTCGTGAGTCTTCGGGTTAACCCCTGCGTTAC	880
Qy	817	GATPACGAATATTTATTTAATACAGTAGCTTCTAAGTCTAAAGATGTTCAATTGAAAAAT	876
Db	881	GACACGGAATATTACCTGATCCCGGAGCTTCTAAGCTCTPAAAGCGTTACACTGTAANAAC	940
Qy	877	ATAACGATTTATATGATTTGACAAATGGCCCATCGTATCTAACGGAAAAATGAATTA	936
Db	941	ATCAGCTAGCTACTATGTACTGACGACCAACGGCCGCTCTACACTAAGGTTAACTAATC	1000
Qy	937	TATTTAGAGGTTATATATATGAGCTPAAATTTATTAATAAAGATTTACACTTAATAT	996
Db	1001	TACTACCGGCTCTGTACAAACGGCCCTGAATTCATCTATCAACGCTACCTCCGACAC	1060
Qy	997	GAAATAGATTTCTTTGGTTAAATCAGGTATTTATTAATTAATATATGATATCATATACAAAT	1056
Db	1061	GAAATCGATTTCTTTGGTTAAATCTGTGTACTTATCAAACTGTACGTTCTTCTTACAAAC	1120
Qy	1057	AATGACACATTTGAGTTATCCGAAAGATGGAATGCCTTTATATATCTTGAATGATTT	1116
Db	1121	AACGACACATGTTGGTTTACCGGAAGAGGTAACGCTTTCAAACAACCTGGACAGATTT	1180
Qy	1117	CTAAGATAGGTTATTAATGCCCAAGTATCCCTCTTTATTAATAAATGGAACAGTAAAA	1176
Db	1181	CTGCGTGGTTGTTACAACGCTCGGGTATCCCGCTGTACAAAAAATGGAAGCGTTTAAA	1240
Qy	1177	TTGCGTGAATTTAAACACGATTTCTGTACAACTTAAATTTATGATGATTAATAATGCACTT	1236
Db	1241	CTGCGTGAACGTAAACCTACGCTGTTCAGCTGAACTGTAGACAGCAAAAAACGCTTCT	1300
Qy	1237	TTAGACATAGTAGTACCATATATGTCAATAATAGCAACGATCCAATATAGGATATATTA	1296
Db	1301	CTGGTCTGGTTGTGATCCACAACGCTGAGATCGTATAGACCCGGAACCGTGCATCTCTG	1360
Qy	1297	ATTGCAACCACTGTTACTTTATCATTTAAAAAGATAAATTTTAGAGTGTGATGGTAC	1356

OS Synthetic.
 XX Key Location/Qualifiers
 FH CDS 108..1529
 FT /*tag= a
 XX
 XX W09808540-A1.
 XX
 XX PD 05-MAR-1998.
 XX
 XX PF 28-AUG-1997; 97WO-US15394.
 XX
 XX PR 28-AUG-1996; 96US-0704159.
 XX
 XX PA (OPHI-) OPHIDIAN PHARM INC.
 XX
 XX PI Thalley BS, Williams JA.
 XX
 XX DR WPI: 1998-230234/20.
 XX P-PSDB: AAW68400.
 XX
 XX PT Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 PT and other animals at risk of intoxication with clostridial toxin
 XX
 XX PS Example 49; Page 376-378; 428pp; English.
 XX
 CC This is the DNA sequence of the Clostridium botulinum serotype G
 CC (113/30 strain) neurotoxin fragment C gene contained in plasmid
 CC pETH15b. The encoded Bo/C fragment C polypeptide (see AAW68400) has a
 CC His-tagged N-terminal extension. The vector can be used to express
 CC native (i.e., non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins, especially type B and type E toxins. Methods
 CC are provided which allow for the isolation of soluble recombinant
 CC proteins free of significant endotoxin contamination. Preferred
 CC hosts for production of recombinant proteins are E. coli, insect
 CC CC immunogens for the production of vaccines and antitoxins that are
 CC useful in the treatment of humans and animals at risk of
 CC intoxication with clostridial toxin.
 CC
 XX Sequence 1535 BP; 596 A; 172 C; 246 G; 521 T; 0 other;
 XX
 SQ
 Query Match 17.8%; Score 248.2; DB 19; Length 1535;
 Best Local Similarity 55.9%; Pred. No. 5.5e-34;
 Matches 547; Conservative 0; Mismatches 408; Indels 24; Gaps 3;

DB 566 AATAATATGACATTAAAGATGTTAAATGCAAAATCTAAATATATTTTTCGAATATAG 625
 OY 501 ACCGATTAATATTTAAAGCTTATTTAGCAAAATGAAGGGTTTAACTACTATCTAATGA 560
 DB 626 TATTAAGATATATATATAGATTTATATATATATATATATATATATATATATATAT 665
 OY 561 TAGATTAATCTCTGCTAAATTTGTATATAAATGAGTACTTAAAGGAGTGCAGAAATTAC 620
 DB 686 TAGATTAATGCAAT 745
 OY 621 TGGTTAGAGCTATTAAGAGGATATATATATATATATATATATATATATATATATAT 680
 DB 746 AAACCTTAGATGAATTAATTTCTAGTAATGATATAGACTTCAAAATTAATTTGTAACA 805
 OY 681 TAAATATCAATACCTTTCTGATTAATATTTAGATATTTTGCAGCAATTAATCCAA 740
 DB 806 TACTACTAATATTTGTTGGATTAAGATTTTATATATATTTTGGTAGAGANTTAATGCTAC 865
 OY 741 AGAGATTGAAAATTAATACACAGTTATTTATCTATATACCTTTTAAAGACTTTGGGG 800
 DB 866 AGAGTATCTCTACTATATATTTGATTCATCAATCACTCAAAATACTTTAAAGATTTTGGGG 925
 OY 801 AAACCTTACGATATGATACGAATATTTATTAATACAGTACCTTACTCTAAGA 860
 DB 926 GAATCCCTTAATAGATGACATACAAATCTATCTGTTAATCAAG-----GTATGCAAAA 979
 OY 861 TGTTCATTTGAAAATTAATACAGATTAATGATTTATTTGCAAAATGCCCATCTATCTAA 920
 DB 960 TATCTATATTAAGATTTTATAGTAAGCTCTATGGGGAAACGACACAGCTCAAACTT 1039
 OY 921 CGGAAATTAATATATATATATATAGAGGTTATATATATATATATATATATATATAT 980
 DB 1040 TAAATATGACGAATTAATATATATATATATATATATATATATATATATATATAT 1099
 OY 961 ATATACCTTAATATGAATGATTTCTTTGTTAATACAGGATTTTATTAATATATA 1040
 DB 1100 AGCATCAATTTCTCGAAATATTAATATATATATATATATATATATATATATATATA 1159
 OY 1041 TGTATCATTAATCAATTAAT 1059
 DB 1160 TCTTATATATGATTAATAT 1178
 RESULT 14
 AAV30581
 ID AAV30581 standard; DNA; 1547 BP.
 XX
 AC AAV30581;
 DT 07-DEC-1998 (first entry)
 XX
 DE Clostridium botulinum toxin B fragment C gene in pETH15b.
 XX
 KW Antitoxin; vaccine; neurotoxin; toxin B; intoxication; immunogen;
 KW botulism; Bo/B; ds.
 XX
 OS Clostridium botulinum serotype B Danish strain.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 108..1526
 FT /*tag= a
 XX
 XX PN W09808540-A1.
 XX PD 05-MAR-1998.
 XX
 XX PF 28-AUG-1997; 97WO-US15394.
 XX PR 28-AUG-1996; 96US-0704159.
 XX
 XX PA (OPHI-) OPHIDIAN PHARM INC.

XX Thalleys BS, Williams JA;
 PI WPI: 1998-230234/20.
 DR P-PSDB; AAM68394.
 XX
 PT Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 PT and other animals at risk of intoxication with clostridial toxin
 XX
 PS Example 35; Page 303-305; 428pp; English.

This is the DNA sequence of the Clostridium botulinum serotype B
 CC (Danish strain) toxin fragment C gene contained in plasmid pETH15b.
 CC The encoded fragment C polypeptide (see AAM68394) has a His-tagged
 CC N-terminal extension. The vector was used to express native
 CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.

Sequence 1547 BP; 634 A; 148 C; 263 G; 502 T; 0 other;

Query Match 16.9%; Score 234.8; DB 19; Length 1547;
 Best Local Similarity 52.4%; Pred. No. 1.1e-31;
 Matches 689; Conservative 0; Mismatches 577; Indels 48; Gaps 6;

90 AAGAGACATTTTAAATTTAGATATTAATGATATTAATCAGATATATCTGGCTT 149
 233 AATTAATATTTATCTTAATTAAGATATAAGATATAATTAATTAATGATATCAGATA 292
 150 TAATCATCTGATATAACATATCCAGATGCTCAATGCTGGCCGCAATTAATGCAAGC 209
 293 TGGGCAAGAGTAGAGATATATGATGAGTCGACCTTAATGATTAATTAATTTAATT 352
 210 AATACATTTAGTAAACAAATGAACTCTGCAAGTTATAGTCATTAAGCTATGATTTGA 269
 353 AACTAGTTAGCAAAA-----TAGTAAGATTAGAGTACCTCAAAATCGAATATCAT 403
 270 AATATATGATTTGTTTAAATTTTACCTAGCTTTGGTGGGTTCCCTAAAGATATC 329
 404 ATTATATAGTGTCTCTCTGATTTTACCTGTTAGCTTTTGGATTAAGATTAATATTA 463
 330 TGTCTAGCTATTAGAACAAATATGGCAAAATAGATTTCAATTAATTTAGCTATGAAA 389
 464 GATGATGATGATACAAATTAATATTCATATGATTAATACATTAATTTATGATGAAA 523
 390 ACATAGCTATCATATGATGATGCTGGTGGAGTATACCTTTAAAGATTAATTAATATG 449
 524 TAAT-----TCGGCTGGGAAATAATCTATTAAGGGATTAATAGATTAATATG 568
 450 GACTTTAAAGATTTCCGGGAGAGACTTAGACAATAATACCTTTAGAGATTACTGTTAA 509
 569 GACTTTAAATTAATTAATTAAGCAAAATCCAAATGCTATTTTGAATATATACATTAAGA 628
 510 ATTATATGCTTTTATAGCAATTAATATGGTGTATTAATACATTAATTAATGATATATG 569
 629 AGATATATCAGATATATTAATATAGATGTTTGTGTAATATTAATTAATATATAT---TTGAA 665
 570 TTCTCTATTTGTTATTAATTAATGAGTACTTTATGGAAGTGCAGAAATTAATGTTAGG 629
 686 TAACCTTAATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 745
 630 AGCATTTAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 689
 746 AGAGTTTATGCTATATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 805
 690 ATAGCTTTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 749

DB 806 ATTTATTTGGTGAATATTTTACATATTTTAAATACGAATTAATGATCAATTAATTA 865
 OY 750 AATATTTACACAGATTAATTTATCTATTAACCTTTTAAAGAGCTCTGGGAAACCTTT 809
 DB 866 AGAAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 925
 OY 810 ACATATGATATGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 869
 DB 926 AATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 985
 OY 870 GAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 929
 DB 986 AAGCAAGATTTACCTCTAGTGAATTTTAAACGATGCAATTAATTAATTAATTA 1045
 OY 930 GAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 987
 DB 1046 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1105
 OY 988 -CCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1046
 DB 1106 TTCTCATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1165
 OY 1047 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1100
 DB 1166 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1225
 OY 1101 TAATCTGATATGATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1160
 DB 1226 ATTTGTTTACCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1285
 OY 1161 AATGAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1220
 DB 1286 ATATGATGAAACCAACATATAGTTGCTGCTTTTAAAGATTAATTAATTAATTA 1345
 OY 1221 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1280
 DB 1346 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1405
 OY 1281 AATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1329
 DB 1406 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1465
 OY 1330 -GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1382
 DB 1466 TAATTTAAATTTGGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1519

RESULT 15
 AAV30579
 ID AAV30579 standard; DNA; 3876 BP.
 AC AAV30579;
 XX
 DT 07-DEC-1998 (first entry)
 XX
 DE Clostridium botulinum type B toxin gene from Danish strain.
 XX
 KW Antitoxin; vaccine; neurotoxin; toxin B; intoxication; immunogen;
 KM botulism; BoTB; ds.
 OS Clostridium botulinum serotype B Danish strain.
 XX
 PN W09808540-A1.
 XX
 PD 05-MAR-1998.
 XX
 PF 28-AUG-1997; 97WO-US15394.
 XX
 PR 28-AUG-1996; 96US-0704159.
 XX
 PA (OPHI-) OPHIDIAN PHARM INC.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:58:27 ; Search time 56 seconds
(Without alignments)
7623.108 Million cell updates/sec

Title: US-09-816-467-3

Perfect score: 1392
Sequence: 1 atggcttttcaacacccaat.....aggatgacaaatgatata 1392

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1379.8	99.1	1858	1 US-08-668-381A-6	Sequence 6, Appl
2	1346.4	96.7	1359	1 US-07-618-312A-1	Sequence 1, Appl
3	1346.4	96.7	1359	1 US-08-280-228-1	Sequence 1, Appl
4	1287.2	92.5	1359	1 US-08-110-786A-7	Sequence 7, Appl
5	799.2	57.4	1359	1 US-07-618-312A-3	Sequence 3, Appl
6	799.2	57.4	1359	1 US-08-280-228-3	Sequence 3, Appl
7	228.2	16.4	3891	1 US-08-480-604A-27	Sequence 27, Appl
8	228.2	16.4	3891	1 US-08-405-496A-27	Sequence 27, Appl
9	228.2	16.4	3891	1 US-08-915-136-27	Sequence 27, Appl
10	220.8	15.9	3509	4 US-09-255-829-19	Sequence 19, Appl
11	83.4	6.0	161	1 US-08-110-786A-6	Sequence 6, Appl
12	80	5.7	161	1 US-08-110-786A-4	Sequence 4, Appl
13	75.6	5.4	1511	1 US-07-991-867B-8	Sequence 8, Appl
14	75.6	5.4	1511	1 US-08-107-755A-8	Sequence 8, Appl
15	75.6	5.4	1511	1 US-08-544-332-8	Sequence 8, Appl
16	75.6	5.4	1511	1 US-09-370-861A-8	Sequence 8, Appl
17	74.2	5.3	1330	1 US-08-480-604A-22	Sequence 22, Appl
18	74.2	5.3	1330	1 US-08-405-496A-22	Sequence 22, Appl
19	74.2	5.3	1330	1 US-08-915-136-22	Sequence 22, Appl
20	74.2	5.3	1402	1 US-08-480-604A-25	Sequence 25, Appl
21	74.2	5.3	1402	1 US-08-405-496A-25	Sequence 25, Appl
22	74.2	5.3	1402	1 US-07-991-867B-32	Sequence 32, Appl
23	73.8	5.3	660	1 US-08-107-755A-32	Sequence 32, Appl
24	73.8	5.3	660	1 US-08-544-332-32	Sequence 32, Appl
25	73.8	5.3	660	4 US-09-370-861A-32	Sequence 32, Appl
26	73.8	5.3	660	4 US-08-852-629-11	Sequence 11, Appl
27	73.8	5.3	4810	3 US-08-852-629-11	Sequence 11, Appl

28	73.8	5.3	4838	3 US-08-852-629-15	Sequence 15, Appl
29	65.4	4.7	19124	2 US-08-487-826B-13	Sequence 13, Appl
30	62.4	4.5	3701	4 US-08-845-258-10	Sequence 10, Appl
31	62.4	4.5	3701	4 US-08-990-571-10	Sequence 10, Appl
32	62.4	4.5	3701	4 US-08-723-142A-10	Sequence 10, Appl
33	62.4	4.5	3701	4 US-09-528-784A-10	Sequence 10, Appl
34	61.8	4.4	19124	2 US-08-487-826B-13	Sequence 13, Appl
35	60.6	4.4	854	4 US-08-998-416-186	Sequence 534, App
36	60.6	4.4	854	4 US-08-998-416-186	Sequence 186, App
37	60.4	4.3	663	4 US-08-998-416-186	Sequence 186, App
38	59.8	4.3	663	4 US-08-998-416-186	Sequence 186, App
39	59.2	4.3	168575	4 US-09-426-290-1	Sequence 191, App
40	58.8	4.2	5433	3 US-08-929-329-1	Sequence 1, Appl
41	58.6	4.2	1323	1 US-08-307-499-36	Sequence 36, Appl
42	58.6	4.2	1323	1 US-09-299-268-36	Sequence 36, Appl
43	58.6	4.2	14176	1 US-08-307-499-1	Sequence 1, Appl
44	58.6	4.2	14176	1 US-08-307-499-14	Sequence 14, Appl
45	58.6	4.2	14176	4 US-09-299-268-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-668-381A-6
; Sequence 6, Application US/08668381A
; Patent No. 5780024
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert H.
; APPLICANT: Fishman, Paul S.
; APPLICANT: Francis, Jonathan W.
; APPLICANT: Hosler, Betsy A.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/263001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1858 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-668-381A-6
Query Match 99.1%; Score 1379.8; DB 1; Length 1858;
Best Local Similarity 99.9%; Pred. No. 7.6e-248;

Matches 1381: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 TCACACCAATTCATTTCTTATTCCTAAAAATCTGATTTGGGTGATTAATGAAGA 69
 |||||
 Db 476 TCACACCACTTCATTTCTTATTCCTAAAAATCTGATTTGGGTGATTAATGAAGA 535

QY 70 GATATGATTTTATTAAGAGACATTTTAAATTTAGATATTAATGATTT 129
 |||||
 Db 536 GATATGATTTTATTAAGAGACATTTTAAATTTAGATATTAATGATTT 595

QY 130 ATATGAGATATTCCTGGCTTAAATTCATCTGTAATAACATATCCAGATGCTCAATGGTG 189
 |||||
 Db 596 ATATGAGATATTCCTGGCTTAAATTCATCTGTAATAACATATCCAGATGCTCAATGGTG 655

QY 190 CCCGGAATTAATGGCAAGCAATACATTTTGTAAACATATCTTGAAGTTATAGTG 249
 |||||
 Db 656 CCCGGAATTAATGGCAAGCAATACATTTTGTAAACATATCTTGAAGTTATAGTG 715

QY 250 CATTAAGCTATGATTTGATATTAATGATATGTTTAAATTTTACCGTTAGCTTTGG 309
 |||||
 Db 716 CATTAAGCTATGATTTGATATTAATGATATGTTTAAATTTTACCGTTAGCTTTGG 775

QY 310 TTGAGGCTTCCTAAAGATATCTGCTAGTATTTAGACAAATATGCGACAATGAGATTTCA 369
 |||||
 Db 776 TTGAGGCTTCCTAAAGATATCTGCTAGTATTTAGACAAATATGCGACAATGAGATTTCA 835

QY 370 ATATTTAGCTCTATGAAAAAACAATGCTATCATAGATCTGGTGGAGTCTACTT 429
 |||||
 Db 836 ATATTTAGCTCTATGAAAAAACAATGCTATCATAGATCTGGTGGAGTCTACTT 895

QY 430 AAAGTAAATTAATTAATGAGATTTTAAAGATTCGCGGAGAGAGTTAGACAATTAAT 489
 |||||
 Db 896 AAAGTAAATTAATTAATGAGATTTTAAAGATTCGCGGAGAGAGTTAGACAATTAAT 955

QY 490 TTGAGGATTTACCTGATTAATTAATGCTTATTAGCAATTAATGAGTTTATTAAT 549
 |||||
 Db 956 TTGAGGATTTACCTGATTAATTAATGCTTATTAGCAATTAATGAGTTTATTAAT 1015

QY 550 ATTAATGATGATTAATTAATGATTTTCTGCTAATTTGATATTAATGAGATCTTATGGAGT 609
 |||||
 Db 1016 ATTAATGATGATTAATTAATGATTTTCTGCTAATTTGATATTAATGAGATCTTATGGAGT 1075

QY 610 GCAGAAATTTACTGTTTAGAGCTATTAGAGAGATTAATTAATTAATTAATTAATTAAT 669
 |||||
 Db 1076 GCAGAAATTTACTGTTTAGAGCTATTAGAGAGATTAATTAATTAATTAATTAATTAAT 1135

QY 670 AGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 729
 |||||
 Db 1136 AGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1195

QY 730 TTAATCCAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 789
 |||||
 Db 1196 TTAATCCAAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1255

QY 790 GACTTCTGGGAAACCTTTACGATATGATACAGAAATTTATTTATTAATCCAGTACTTCT 849
 |||||
 Db 1256 GACTTCTGGGAAACCTTTACGATATGATACAGAAATTTATTTATTAATCCAGTACTTCT 1315

QY 850 AGTTCTAAGAGATTTCAATTTGAAAAATTAACAGATTTATTAATTAATTAATTAATTAATTAAT 909
 |||||
 Db 1316 AGTTCTAAGAGATTTCAATTTGAAAAATTAACAGATTTATTAATTAATTAATTAATTAATTAAT 1375

QY 910 TCGTATTAACGAGAAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 969
 |||||
 Db 1376 TCGTATTAACGAGAAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1435

QY 970 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1029
 |||||
 Db 1436 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1495

QY 1030 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1089
 |||||
 Db 1496 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1555

QY 1090 AATGCCCTTATATATCTTGATAGATTTCTAAGATAGGTTATTAATGCCCGCAGGTATCCCT 1149
 |||||
 Db 1556 AATGCCCTTATATATCTTGATAGATTTCTAAGATAGGTTATTAATGCCCGCAGGTATCCCT 1615

QY 1150 CTTTATTAATAAATAAATGAAGCAGTAAATTCGCTGATTTTAAAAACCTATTTCTGTCACACTT 1209
 |||||
 Db 1616 CTTTATTAATAAATAAATGAAGCAGTAAATTCGCTGATTTTAAAAACCTATTTCTGTCACACTT 1675

QY 1210 AATTAATATGATTAATAAATGATCTTTAGGACTAGTAGTACCCTTAATGGTCAATA 1269
 |||||
 Db 1676 AATTAATATGATTAATAAATGATCTTTAGGACTAGTAGTACCCTTAATGGTCAATA 1735

QY 1270 GGCACAGATCCAAATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1329
 |||||
 Db 1736 GGCACAGATCCAAATAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1795

QY 1330 GATTAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
 |||||
 Db 1796 GATTAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1855

QY 1390 TAA 1392
 |||||
 Db 1856 TAA 1858

RESULT 2
 US-07-618-312A-1
 ; Sequence 1, Application US/07618312A
 ; Patent No. 5389540
 ; GENERAL INFORMATION:
 ; APPLICANT: Makoff Dr, Andrew J
 ; APPLICANT: Romanos Dr, Michael A
 ; APPLICANT: Clare Dr, Jeffrey J
 ; APPLICANT: Fairweather Dr, Neil F
 ; TITLE OF INVENTION: VACCINES
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: 14th Floor
 ; STREET: 2200 Clarendon Boulevard,
 ; CITY: Arlington,
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22201
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/618, 312A
 ; FILING DATE: 19910516
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 8926832.0
 ; FILING DATE: 28-NOV-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9006097.1
 ; FILING DATE: 17-MAR-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crawford Mr, Arthur R
 ; REGISTRATION NUMBER: 25,327
 ; REFERENCE/DOCKET NUMBER: 510-51
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 0101 703 8750400
 ; TELEFAX: 0101 703 5253468
 ; TELEX: 200797 NIXN UR
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1359 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear

```

? MOLECULE TYPE: DNA (genomic)
?
? HYPOTHETICAL: NO
?
? ORIGINAL SOURCE:
?
? ORGANISM: Clostridium tetani
?
? FEATURE:
?
? NAME/KEY: CDS
?
? LOCATION: 1..1356
?
US-07-618-312A-1

```

Query Match	96.7%;	Score 1346.4;	DB 1;	Length 1359;
Best Local Similarity	99.6%;	Pred. No. 1.2e-241;		
Matches 1350;	Conservative	0;	Mismatches 6;	Indels 0;
				Gaps 0

QY	37	AAAAATCTGGATGTGGGCTGTATTAATGAAGAAGATATGATGTTATATATAAAAAAGACT	96
Db	4	AAAATCTGGATGTGGGCTGTATATATGAAGAAGATATGATGTTATATATAAAAAAGACT	63
QY	97	ACAATTTTAAATTTAGATTTTAAATATGATATTTAATATCGATATATCTGGGTTTAAATCA	15
Db	64	ACAATTTTAAATTTAGATTTTAAATATGATATTTAATATCGATATATCTGGGTTTAAATCA	12
QY	157	TCMGTAATACATATCCAGATGCTCAATTTGGTCCCGGATTAATGCGAAGCAATACAT	21
Db	124	TCTGTATATACATATCCAGATGCTCAATTTGGTCCCGGATTAATGCGAAGCAATACAT	18
QY	217	TTAGTAACAATGAATCTCTGTAAGTTATAGTGCATPAAGCATATGATATTAATATAT	27
Db	184	TTAGTAACAATGAATCTCTGTAAGTTATAGTGCATPAAGCATATGATATTAATATAT	24
QY	277	GATATGTTTAATATATTTACCGTTAGCTTTGGTGTAGAGGGTCTTAAAGTATCTGCTAGT	33
Db	244	GATATGTTTAATATATTTACCGTTAGCTTTGGTGTAGAGGGTCTTAAAGTATCTGCTAGT	30
QY	337	CATTAGAACAAATATGGCAAAATAGTATTCATTAATTTAGCTGTATGAATAAAACATAGT	39
Db	304	CATTAGAACAAATATGGCAAAATAGTATTCATTAATTTAGCTGTATGAATAAAACATAGT	36
QY	397	CTATCAATAGCATCTGGTTGAGTGTATCACTTAAAGTAAATTAATCTTAAATATGCACTTTA	45
Db	364	CTATCAATAGCATCTGGTTGAGTGTATCACTTAAAGTAAATTAATCTTAAATATGCACTTTA	42
QY	457	AAAGATTCGGGGGAGAGTTAGACAAATTAACCTTTGGGATTAACCTGATTAATTTAAT	51
Db	424	AAAGATTCGGGGGAGAGTTAGACAAATTAACCTTTGGGATTAACCTGATTAATTTAAT	48
QY	517	GCTTATTTAGCAAAATTAATGGGTTTTTATAAATCTTACTATGATAGATTAATCTTCTGCT	57
Db	484	GCTTATTTAGCAAAATTAATGGGTTTTTATAAATCTTACTATGATAGATTAATCTTCTGCT	54
QY	577	AATTTGATATTAATTAATGGAGTCTATATGGGAAGTGCACAAATTAATCTGGTTTAGAGCTAAT	63
Db	544	AATTTGATATTAATTAATGGAGTCTATATGGGAAGTGCACAAATTAATCTGGTTTAGAGCTAAT	60
QY	637	AGAGAGCATTAATTAATPACATTTAAACTAGATGATATATAATTAATATCATACAGTT	69
Db	604	AGAGAGCATTAATTAATPACATTTAAACTAGATGATATATAATTAATATCATACAGTT	66
QY	697	TCATTTGATTAATTTAGGATATTTGCAAGCATTAATCCAAAAGAGATTGAATAATTA	75
Db	664	TCATTTGATTAATTTAGGATATTTGCAAGCATTAATCCAAAAGAGATTGAATAATTA	72
QY	757	TACCAAGATTATTTATCTATPACCTTTTAAAGAGACTCTGGGGAAACCTTTACGATAT	81
Db	724	TACCAAGATTATTTATCTATPACCTTTTAAAGAGACTCTGGGGAAACCTTTACGATAT	78
QY	817	GATTCACAATATTAATTAATACCAAGTATGCTTACGTTCTAAGCATGTTCAATTTGAAAAAT	87
Db	784	GATTCACAATATTAATTAATACCAAGTATGCTTACGTTCTAAGCATGTTCAATTTGAAAAAT	84
QY	877	ATAACAGATTATATGATTTTGACAATATGCGCATCTGATATCTAAGCGAAAAATTTGAATAT	93
Db	844	ATAACAGATTATATGATTTTGACAATATGCGCATCTGATATCTAAGCGAAAAATTTGAATAT	90

QY	937	TATTATAGAGGCTTATTAATGAGCTAAATTTTATATATAAAGATATACACTATATAT	996
Db	904	TATTATAGAAAGGTTATATATATGACTAAATTTTATATATAAAGATATACACTATATAT	963
QY	997	GAATATGATCTTTTGGTTAATCAGTGATTTTATTAATATATATATGATCATATACAAAT	1056
Db	964	GAATATGATCTTTTGGTTAATCAGTGATTTTATTAATATATATGATCATATACAAAT	1022
QY	1057	AATGACACACTTGTAGTGTATCCGAAAGATGGAATGCCCTTTAATTAATCTTGATGAAAT	1116
Db	1024	AATGACACCAATGTAGTGTATCCGAAAGATGGAATGCCCTTTAATTAATCTTGATGAAAT	1083
QY	1117	CTAAGAGTAGTGTATATGCGCCAGGATGCCCTTTATATATAAANAATGGAAGCAGTAAAT	1176
Db	1084	CTAAGAGTAGTGTATATATGCGCCAGGATGCCCTTTATATATAAANAATGGAAGCAGTAAAT	1143
QY	1177	TTCGCTGATTTAAANAACCTATCTGTACCACTTAATATATGATGATATAAANAATGCACTCT	1236
Db	1144	TTCGCTGATTTAAANAACCTATCTGTACCACTTAATATATGATGATATAAANAATGCACTCT	1209
QY	1237	TTAGGACCTGTAGGTACCCATTAATGGCAATATGGCAAGCATCCCAATTAGGGATATATTA	1296
Db	1204	TTAGGACCTGTAGGTACCCATTAATGGCAATATGGCAAGCATCCCAATTAGGGATATATTA	1263
QY	1297	ATTGCAAGCACTGTGACTTTATCATTTTAAAGATATAATTTTGAAGATGATGTGGTAC	1356
Db	1264	ATTGCAAGCACTGTGACTTTATCATTTTAAAGATATAATTTTGAAGATGATGTGGTAC	1323
QY	1357	TTTGTACCTACAGATGAGGAGTGGACCAATGATTTAA	1392
Db	1324	TTTGTACCTACAGATGAGGAGTGGACCAATGATTTAA	1359

RESULT 3
 US-08-280-228-1
 Sequence 1, Application US/08280228
 Patent No. 5571694
 GENERAL INFORMATION:
 APPLICANT: Makoff Dr, Andrew J
 APPLICANT: Romanos Dr, Michael A
 APPLICANT: Claire Dr, Jeffrey J
 APPLICANT: Fairweather Dr, Neil F
 TITLE OF INVENTION: VACCINES
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHAYE P.C.
 STREET: 1100 No. 5571694th Glebe Road
 City: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/280,228
 FILING DATE: 25-JUL-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/618,312
 FILING DATE: 27-NOV-1990
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 8926832.0
 FILING DATE: 28-NOV-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9006097.1
 FILING DATE: 17-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Wilson, Mary J.
 REGISTRATION NUMBER: 32,955

```

REFERENCE/DOCKET NUMBER: 117-163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
type: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-08-280-228-1

```

Query Match 96.7%; Score 1346.4; DB 1: Length 1359;

Best Local Similarity 99.6%; Pred. No. 1.2e-241; Mismatches 6; Indels 0; Gaps 0;

```

Matches 1350; Conservative 0;
QY 37 AAAAATCGATGTTGGTTGATATGATGAAGATATGATGTTATATTAATAAAGAGT 96
DB 4 AAAATCTGATGTTGGTTGATATGATGAAGATATGATGTTATATTAATAAAGAGT 63
QY 97 ACAATTTAAATTTAGATATTAATTAATGATATTAATGATATATATCTGTTTAAATCA 156
DB 64 ACAATTTAAATTTAGATATTAATTAATGATATTAATGATATATATCTGTTTAAATCA 123
QY 157 TCTGTATTAATCAATTCAGATGCTCAATGCTGCGCGGAATTAATGCAAAAGCAATTCAT 216
DB 124 TCTGTATTAATCAATTCAGATGCTCAATGCTGCGCGGAATTAATGCAAAAGCAATTCAT 183
QY 217 TTACTAACAATGATGATCTGGAATTAATGATGATGATGATGATGATGATGATGATGAT 276
DB 184 TTACTAACAATGATGATCTGGAATTAATGATGATGATGATGATGATGATGATGATGAT 243
QY 277 GATATGTTTAAATTTTACCGTTAGCTTTTGGTTGAGGGTCTCTAAAGTATCTGCTAGT 336
DB 244 GATATGTTTAAATTTTACCGTTAGCTTTTGGTTGAGGGTCTCTAAAGTATCTGCTAGT 303
QY 337 CATTTAGAACATATGCGACAATGAGTATCATTAATTTAGCTCTATGAATAAATAGT 396
DB 304 CATTTAGAACATATGCGACAATGAGTATCATTAATTTAGCTCTATGAATAAATAGT 363
QY 397 CTATCAATAGATCTGTTGGATGATCTACTTAAGGTAATTAATTAATTAATTAATTAAT 456
DB 364 CTATCAATAGATCTGTTGGATGATCTACTTAAGGTAATTAATTAATTAATTAATTAAT 423
QY 457 AAAGATTCGCGGAGAGATTAAGACAATACTTTTGGGATTAACCTGATTAATTAATTAAT 516
DB 424 AAAGATTCGCGGAGAGATTAAGACAATACTTTTGGGATTAACCTGATTAATTAATTAAT 483
QY 517 GCTTATTTAGCAATAATGCGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 576
DB 484 GCTTATTTAGCAATAATGCGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 543
QY 577 AATTTGATTAATTAATGATCTTAATGGAAGTGCAGAAATTAATTAATTAATTAATTAAT 636
DB 544 AATTTGATTAATTAATGATCTTAATGGAAGTGCAGAAATTAATTAATTAATTAATTAAT 603
QY 637 AGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
DB 604 AGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
QY 697 TCTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
DB 664 TCTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
QY 757 TAGCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816

```

```

DB 724 TAGCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
QY 817 GATPACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
DB 784 GATPACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY 877 ATPACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
DB 844 ATPACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
QY 937 TATTTAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
DB 904 TATTTAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
QY 997 GAATATGATCTTTGTTAAATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
DB 964 GAATATGATCTTTGTTAAATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
QY 1057 ATGAGCACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116
DB 1024 ATGAGCACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083
QY 1117 CTAGAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
DB 1084 CTAGAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
QY 1177 TTGGGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1236
DB 1144 TTGGGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
QY 1237 TTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
DB 1204 TTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
QY 1297 ATTCAGAGCACTGCTACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
DB 1264 ATTCAGAGCACTGCTACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1357 TTTGTACCTACAGATGAGGATGAGCAATGATTA 1392
DB 1324 TTTGTACCTACAGATGAGGATGAGCAATGATTA 1359

RESULT 4
US-08-110-786A-7
Sequence 7, Application US/08110786A
Patent No. 5443966
GENERAL INFORMATION:
APPLICANT: FAIRWEATHER, Neil Fraser
APPLICANT: MAKOFF, Andrew Joseph
TITLE OF INVENTION: Expression of tetanus toxin fragment C
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: Nixon & Vanderhye P.C.
STREET: 1100 No. 5443966th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110,786A
FILING DATE: 23-AUG-1993 1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/777,337
FILING DATE: 29-NOV-1991
PRIOR APPLICATION DATA:

```


|||||

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/618,312A
FILING DATE: 19910516
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
APPLICATION NUMBER: GB 9006097.1
ATTORNEY/AGENT INFORMATION:
NAME: Crawford Mc, Arthur R
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 510-51
TELECOMMUNICATION INFORMATION:
TELEPHONE: 0101 703 8750400
FAX: 0101 703 5253468
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-07-618-312A-3

Query March 57.4%; Score 799.2; DB 1; Length 1359;
Best Local Similarity 74.3%; Pred. No. 2.9e-140;
Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

```

```

Db 484 GCCTACCTGGCTAACAAATGGTTTTCATCCTATCATCAACAGCTGCTCTTCTGCT 543
Qy 577 AATTGTATTAATAGAGTACTTATGGGAAGTGCAGAAATATCTGGTTAGAGCTATT 636
Db 544 AACCTGTACATCAACGCGTTCTGATGGCTCCGCTGAATTCACCTGCTGGGGCTATC 603
Qy 637 AGAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
Db 604 CGTGAGGACAAACATCATCTTAAGCTGACCGTTGCAACCAACCAACAGCTAGCTA 663
Qy 697 TCTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
Db 664 TCCATCGCAAGTCCCTATCTCTGCAAGCACTGAACCGGAAGAGATCGAAAACTG 723
Qy 757 TACACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
Db 724 TATACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
Qy 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
Db 784 GACACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
Qy 877 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
Db 844 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
Qy 937 TATATGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
Db 904 TACTACGAGCTCTGTACACAGCGCTGAATTCATCAACAGCTACACCTGACCAAC 963
Qy 997 GAAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
Db 964 GAAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
Qy 1057 AATGACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
Db 1024 AACGACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
Qy 1117 CTAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
Db 1084 CTGCGTGTGTTACACAGCTCCGCGTATCCGCTGTACAAAAATGGAACCTTTTAA 1143
Qy 1177 TTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1236
Db 1144 CTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
Qy 1237 TTAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1296
Db 1204 CTGCGTGTGTTGTTACACAGCTCCGCGTATCCGCTGTACAAAAATGGAACCTTTTAA 1263
Qy 1297 ATTGCAAGCACTGTACTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
Db 1264 TTGCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
Qy 1357 TTGCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1392
Db 1324 TTGCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1359

RESULT 6
US-08-280-228-3
; Sequence 3, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 No. 5571694th Glebe Road

```

```

CITY: Arlington,
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/280,228
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/618,312
FILING DATE: 27-NOV-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-163
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-08-280-228-3

```

Query Match 57.4%; Score 799.2; DB 1; Length 1359;
Best Local Similarity 74.3%; Pred. No. 2,9e-140;
Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

```

37 AAAAACTCGATTGGTGGGTGATATGAAGAAGATATAGATGTTATTTAAAGAGAGT 96
4 AAAAACCTGATTGTTGGTGGTGCACGAAGAGACATGATGTTATCCGAAAAAGTCT 63
97 ACAATTTAAATTTAGATATTAATATGATATATATCATGATATATCTGGGTTAAATTC 156
64 ACCATTCTGAACCTTGACATCAACAACGATATATATCTCGACATCTCTGTTTCAACTCC 123
157 TCTGTAATTAACATATCCGAGATGCTCAATTTGGTCCCGGAATTAATGCAAGCAATACAT 216
124 TCTGTAATTAACATATCCGAGATGCTCAATTTGGTCCCGGAATTAATGCAAGCAATACAT 183
217 TTAGTAACAATAGATCTCTGAGTATAGTCAATTAAGTATGATATATTAATTAAT 276
184 CTGTTAAACAAGATCTCTGATTAATTTATCTGTCACAAAGGCAATGACATCAATCAAC 243
277 GATATGTTAAATTAATTTAGCGTTAGTGGTTCAGGGTTCCTAAAGTATCTGCTAGT 336
244 GACATGTTCAACAACCTTCACCGTTAGCTTTCGGCTCGCGGTTCCGAAATTTCTGCTTCC 303
337 CATTTAGACATATGACCAATGAGTATTCATTAATTAAGCTATGAAAAACATACT 396

```

```

Db 304 CACCTGGACAGTACGGCAGTAAAGAGTACATCCATCAGCTATGAAAGAACTCC 363
QY CTATCAATAGAGTATGTTGGATGATATCACTTAAGGTAATTAATTAATGAGCTTAA 456
Db CTGTCCATCGGCTCTGTTGGTCTGTTTCCGTAAGGTAACACCTGATCTGAGCTG 423
QY AAGATTTCCCGGGGAGAAGTTAGCAAAATTAATTTAGGATTTACCTGATTAATTAAT 516
Db AAAGACTCCCGGGGAGAAGTTAGCAAAATTAATTTAGGATTTACCTGATTAATTAAT 483
QY GCTTATTTAGCAATTAATGAGGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 576
Db GCTTATTTAGCAATTAATGAGGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 484
QY AATTGATTAATTAATGAGTATTTAGGAAAGTGCAGAAATTAATTAATTAATTAATTAAT 636
Db AACCTGATCAATCAACGCGTTGATGAGGCTCCGTAATCACTGCTGCTGCTGCTGCT 603
QY AGAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
Db CGTAGAGACACAACTCACTTTAAGCTGACCGTTGCAACCAACCAACCACTAGCTA 663
QY TCTATTGATAAATTTAGATATTTTGAAGCATTAATTAATTAATTAATTAATTAATTAAT 756
Db TCCATGACAGATTTCCGATCTCTGCAAAAGCACTGAACCGGAAGATGGAAGAACTG 723
QY TACACAGTTATTTATCTATTAACCTTTTAAGAGACTTGGGGAACCCCTTAAGATAT 816
Db TATACAGCTACCTGCTATCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
QY GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
Db GACACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
Db ATCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
QY TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 966
Db TACTACGACGCTGCTGACAGCGCTGGAATTCATCATCAACGCTACCTCGAACAAC 963
QY GAAATAGATTTCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
Db GAAATAGATTTCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 964
QY AATGACATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
Db AATGACATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
QY CTAAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
Db CTGGTGTGTTGTTTACACGCTCCGGGTATCCCGCTGTACAAAAAATGGAAGCTGTTAA 1143
QY TTGGCTGATTTAAAACTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAATTAAT 1236
Db CTGGTGTGTTGTTTACACGCTCCGGGTATCCCGCTGTACAAAAAATGGAAGCTGTTAA 1144
QY TTAGGACTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1296
Db CTGGTGTGTTGTTTACACGCTCCGGGTATCCCGCTGTACAAAAAATGGAAGCTGTTAA 1263
QY ATTGACAGCACTGCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
Db ATTGACAGCACTGCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1264
QY TTGCTACCTTACAGATGAGGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1392
Db TTGCTACCTTACAGATGAGGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1324

```

RESULT 7
US-08-480-604A-27

```

Sequence 2: Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KIRK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3888
US-08-480-604A-27

```

Db	2678	GGTATGATCAAAAATAATATATTGGTACTAAAGTAAATTTTGGATCC---AATAGATTAAAA	2734
Qy	206	AAGCAATACATTTTGTAAACAATGATCTTCGAGTTATAGTCATTAAGCTATAGATA	265
Db	2735	ATCCAAATTCATTTATTATTATTAGAAAGAGTAAATTTAGAGTATTTTAAAAATGGCTA	2794
Qy	266	TTGAATTAATATGATATGTTTATTAATTTTACCCTTAGCTTTTGGTGGGGTTCCTAAG	325
Db	2795	TTGTATATAAATAGTATGATGATAAATTTTGTAGTACGCTTTTGGATAGCAATTTCTTAAGT	2854
Qy	326	TATCGTAGTCAATTTAGACAAATATGGCACAAATGAGATTCCAATTAATTAAGCTGTAGA	385
Db	2855	ATTTTAACAGT-----ATAGCTCTAATTAATGAATATATCAATTAATTAATTTGATGG	2905
Qy	386	AAAAACATAGTCTATCAATAGATGATCGTTGGTAGTGTATCACTTTAAGCTAATTAAGTTAA	445
Db	2906	AAAATAAT-----TCAGATGGGAAAGTATCACTTTATTTATGGTAAATTA	2950
Qy	446	TATGACCTTTAAACATTCGCGGGGGAAGTTAGACAAATTAATCTTTTATGAGATTTACCTG	505
Db	2951	TCGTGACCTTTACAGGATCTCGAGAAATTAACAAGAGTAGTTTAAATACAGTCAAA	3010
Qy	506	ATTAATTTATGCTTATTTAGCAAAATAATGGGTTTTTATTAATTTACTAATGATAGAT	565
Db	3011	TCGATTATATATACAGATTATTAATACAGATGATTTTGTAACTATACATTAATTAATAGAT	3070
Qy	566	TATCTTCTGCTAATTTGTATATAAATGAGAGTCTTATGGAAATGCGAAATTAAGTGGTT	625
Db	3071	TAAATTAACCTAAATTTTATTAATGAGAGATTAAATAGATCAAAAAACCAATTTCCAATT	3130
Qy	626	TGGAGCCTTTATGAGAGAGATTAATTAATTAACATTAAACATAGATAGATTAATTAATA	685
Db	3131	TAGGTATATTTCACTGCTAGATAAATATATATGTTTAAATTAGATGTTGTAGAGATACAC	3190
Qy	686	ATCAATACGTTCTATTTGATAAATTTAGCATTTTTCGCAAGCATTAATCCAAAAAGAGA	745
Db	3191	ATGATATATTTTGGATTAATAATTTTAACTTTTGTATGAGAAATTAATGAAGAAABAAA	3250
Qy	746	TTGAAAAATTAATCACACAGTTATTTATCTTAATCTTTTATAGAGACTTCGCGGGAAC	805
Db	3251	TCAAAGTTTATATGATATCAATCAAAATTCAGGATTTTAAAAAGACTTTTGGGGTGATT	3310
Qy	806	CTTTAGCATATGATACGAATATTTTAAATPACAGTAGCTTCTAGTTCTAAGAGATGTC	865
Db	3311	ATTTACAATATGATAACCATCTACTATATGTTAAATTTATATGATCCAAATTAATATGTCG	3370
Qy	866	AATTTGAAAA-----TATACAGATATATNGTATTTGACAAATGCGCCATCGTATCTA	919
Db	3371	ATGTAATTAATGATGATTTAGAGCTTATATGATCTTTAAAGGCCCTAGAGAGTACGCTAA	3430
Qy	920	ACGGAATATGAATATATATTTATGAGAGGTATATTAATGACATAAATTTATTAATAAA	979
Db	3431	TGACTACAAACATTTATTTAAATTTCAAGTTTGTATAGGGGACAAAATTTATTTAATAAA	3490
Qy	980	GATATACACCTAATATATGAATAGATTCCTTTGTTAATCAAGTGATTTTATTAATATAT	1039
Db	3491	AATATGCTTCGGAAT---AAACATATATATGTTAGAAATTAATGATCGTATATATATTA	3547
Qy	1040	ATGTATCATATTAACATTAATGAGACATTTAGAGT1074	
Db	3548	ATGTAGTAGTTAAAAAATTAACAAATATAGCTTAGCT3582	

Query Match	Similarity	16.4%	Score 228.2	DB 1	Length 3891
Best Local	Similarity	55.9%	Pired. No. 2e-34		
Matches 536	Conservative	0	Mismatches 403	Indels 36	Gaps 5
Qy	86	TAAAGAAGACATACATTTTAATTTACATATTAATAATGATATTATATACAGATATATCTG	145		
Db	2618	TTATTAATACCTCTCATATTGCAATTTAGAAATGAAAGCAATTAATCATTTTAATAGACTTTATCTA	2677		
Qy	146	GCTTAATTCATCTCGTATATACATATCCAGATGCTCAATTCAGTCCCGGAGTAATTAAGCA	205		

```

RESULT 8
US-08-405-496A-27
: Sequence 27. Application US/08405496A
: Patent No. 5919665
:
: GENERAL INFORMATION:
:
: APPLICANT: WILLIAMS, JAMES A.
:
: TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
:
: TITLE OF INVENTION: NEUROTOXIN
:
: NUMBER OF SEQUENCES: 30

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 09:49:42 ; Search time 64 seconds

(Without alignments)
8376.261 Million cell updates/sec

Title: US-09-816-467-3

Perfect score: 1392

Sequence: 1 atggcttttcaacccaat.....aggatgacaaatgatata 1392

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 341543 seqs, 19257720 residues

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCPUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248.2	17.8	3894	12 US-10-051-952-9	Sequence 9, Appl1
2	233.2	16.8	3876	12 US-10-051-952-3	Sequence 3, Appl1
3	231.2	16.6	3876	12 US-10-051-952-2	Sequence 1, Appl1
4	228.2	16.4	3891	12 US-10-051-952-1	Sequence 2, Appl1
5	228.2	16.4	4355	10 US-09-288-326-10	Sequence 10, Appl1
6	215.8	15.5	2452	10 US-09-350-756-7	Sequence 7, Appl1
7	213	15.3	1327	10 US-09-350-756-9	Sequence 7, Appl1
8	208.4	15.0	3759	12 US-10-051-952-7	Sequence 7, Appl1
9	206.2	14.8	1293	8 US-08-981-087A-5	Sequence 5, Appl1
10	205.6	14.8	3755	12 US-10-051-952-6	Sequence 6, Appl1
11	181.4	13.0	3825	12 US-10-051-952-8	Sequence 8, Appl1
12	155.8	11.2	3831	12 US-10-051-952-5	Sequence 5, Appl1
13	145.2	10.4	1987	10 US-09-350-756-8	Sequence 8, Appl1
14	113.8	8.2	3876	12 US-10-051-952-4	Sequence 4, Appl1
15	93.8	6.7	1400	10 US-09-350-756-4	Sequence 4, Appl1
16	74.2	5.3	1338	10 US-09-350-756-1	Sequence 1, Appl1
17	62.4	4.5	3701	9 US-09-286-488-10	Sequence 10, Appl1
18	62.4	4.5	3701	10 US-09-737-178-10	Sequence 10, Appl1
19	62	4.5	1313	8 US-08-981-087A-6	Sequence 6, Appl1

20	61.6	4.4	640681	10 US-09-790-988-1	Sequence 1, Appl1
21	61.4	4.4	702	10 US-09-350-756-10	Sequence 10, Appl1
22	59.4	4.3	640681	10 US-09-790-988-1	Sequence 1, Appl1
23	56.2	4.0	2966	10 US-09-737-178-103	Sequence 103, App
24	56.2	4.0	3011	10 US-09-737-178-112	Sequence 112, App
25	55.8	4.0	302250	10 US-09-962-832-154	Sequence 154, App
26	55.4	4.0	1341	10 US-09-350-756-2	Sequence 2, Appl1
27	54.8	3.9	11474	10 US-09-816-028A-1	Sequence 1, Appl1
28	54	3.9	1347	10 US-09-350-756-6	Sequence 6, Appl1
29	53.6	3.9	94	10 US-09-993-292A-15	Sequence 15, Appl1
30	53.6	3.8	416	10 US-09-960-352-4584	Sequence 4584, Ap
31	52.6	3.8	53332	10 US-09-801-861-3	Sequence 4582, Ap
32	52.6	3.8	3701	9 US-09-286-488-10	Sequence 3, Appl1
33	52.4	3.8	3701	10 US-09-737-178-10	Sequence 10, Appl1
34	52.4	3.8	12003	10 US-09-764-877-3976	Sequence 3976, Ap
35	51.8	3.7	1431	10 US-09-774-414-2	Sequence 2, Appl1
36	51.6	3.7	1230	10 US-09-845-335-1	Sequence 1, Appl1
37	51	3.7	1230	10 US-09-764-877-3976	Sequence 3976, Ap
38	50.8	3.6	12003	10 US-09-845-335-1	Sequence 1, Appl1
39	50.4	3.6	446	10 US-09-960-352-3400	Sequence 3400, Ap
40	50.4	3.6	516	10 US-09-960-352-5785	Sequence 5785, Ap
41	50.2	3.6	4441	10 US-09-796-348-20	Sequence 20, Appl1
42	49.8	3.6	1109	9 US-09-938-842A-3333	Sequence 3333, Ap
43	49.6	3.6	431	10 US-09-960-352-5558	Sequence 5558, Ap
44	49.6	3.6	6152	9 US-09-742-096-1	Sequence 1, Appl1
45	49.4	3.5	5046	10 US-09-725-735A-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-10-051-952-9
Sequence 9, Application US/10051952
Patent No. US20020107199A1
GENERAL INFORMATION:
APPLICANT: Walker, Patricia
TITLE OF INVENTION: Methods of Administering Botulinum Toxin
FILE REFERENCE: 2933CIP
CURRENT APPLICATION NUMBER: US/10/051,952
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 09/730,237
PRIOR FILING DATE: 2000-12-05
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 3894
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: primers used to introduce Stu I and EcoR I
OTHER INFORMATION: restriction sites into the 5' and 3' ends of the
OTHER INFORMATION: BONT/A-L chain gene fragment
US-10-051-952-9

Query Match 17.8% Score 248.2; DB 12; Length 3894;
Best Local Similarity 55.9%; Pred. No. 1.3e-35;
Matches 547; Conservative 0; Mismatches 408; Indels 24; Gaps 3;

QY 81 TATATTAATAAGAGTACATTTTAAATTTAGATATTAATTAATGATATTTATCATAT 140
DB 2589 TAAATTTAGTATGATATGATATTTAGTATTTAGTATTTAGAGTGGCGTTTATGATTTT 2648
QY 141 ATCTGGGTTTAAATTTATCTGTAATACATATCCAGATGCTCAATGCTGCCGGAATAA 200
DB 2649 ATCTGGATATGCTGCACTATGATGATAGTATGATCTTAATGATATGAGAA 2708
QY 201 TGGCAAGCAATACATTTAGTAAACATGATCTTCTGAGTATATGTCATTAAGCTAT 260
DB 2709 TGGTCA---ATTTAATTAATTAATATTTCTGAAATATGTAATTAATTAATTAAGTAA 2765
QY 261 GGATATTAATTAATTAATGATATGTTTAAATTAATTTTACCGTTAGCTTTGGTGGAGGCTTC 320


```

Db 3516 TAATTTAATCAAGAGTGGAGATATATACCTTAATATTTTAAAGAAAGAGCAAGAAA 3575
OY 1101 TAATCTGATACAAATCTAAGAGTAGTTATATATGCCCCAGATACCTCTTTATAAAAA 1160
Db 3576 ATTGTTTACCTCTATTAAGATGATTCATGATGATTTAGAAATCTATACAAATTAAGAA 3635
OY 1161 AATGAACAGTAATATTCGCGATTTAAAAACCTATTCGTACAACTTAATTTATGA 1220
Db 3636 ATATGATCAACAGCCACATATATATGTCAGTTCCTTTTAAAAAAGATGAAGAAAGTAC 3695
OY 1221 TGATAAAAATCACTTTAGAGACTAGTAGTACCATATATGTCAAAATAGCAACGATCC 1280
Db 3696 TGATGAGTAGAGATGATGATGATGATTCATGCTTTTCTAGGAATCTGCAATTTGATGA 3755
OY 1281 AATATGAGATATATTAATTCGAAGCAACTGTACTTTATCATTTAAAA----- 1329
Db 3756 GTATTAAGATATTTTCTATATAGTAAATGCTAGTTAAAGAGTAAGAAAGCAATA 3815
OY 1330 -GATAAATTTTAGAGTGTGTACTCTTGTACTCTACATGAGATGAGGATGAC 1382
Db 3816 TAATTTAAATTTGGAGATGTAATTTGGAGTTATTCCTAAAGATGAAGGTGAGC 3869

```

RESULT 3

```

US-10-051-952-2
; Sequence 2, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3876
; TYPE: DNA
; ORGANISM: botulinum toxin
; US-10-051-952-2

```

Query Match 16.6%; Score 231.2; DB 12; Length 3876;
 Best Local Similarity 52.7%; Pred. No. 1.3e-32;
 Matches 693; Conservative 0; Mismatches 573; Indels 48; Gaps 7;

```

OY 90 AAAGGTCAATTTTAAATTTTGAATTTATATATGATTTATATCATGATATATCTGGTT 149
Db 2583 AAATATATTTATCTTAATTTAAGATATATAGATATATATTAATTTTATGAGATA 2642
OY 150 TAATTCATCTGTAATTAACATATCCAGATGCTCAATTTGGCCCGCAATTAATGCGAAGC 209
Db 2643 TCGACCAAGGTAGAGGTATA-----TGATGGGGTCAAGCTTAATGATTAATAATCA 2693
OY 210 AATACATTTAGTAAACATGAATCTTCTGAACTTATAGTCAATTAAGCATGATGATCA 269
Db 2694 ATTAAATTTAATCACTGACGAGATAGTATAGATATAGCACTCAAAATACAGATATATAT 2753
OY 270 ATATATGATATGTTTAATTAATTTTACGCTTACGCTTGTGAGGTTCTCTAAAGTATC 329
Db 2754 ATTATATGATGTTTCCCTGATTTTACGCTTGTGAGGTTGTGAGGATTAATGATTAATG 2813
OY 330 TGCTAGTATTTAGAACATATATGCGCAATGAGTATTCATTAATTTAGCTGATGAATA 389
Db 2814 GAATCATATATTAACAAATATATTTATTCATATGATATATGATTAATTTATGATGAATA 2873
OY 390 ACATAGCTATACATAGATGATGAGTGTGAGGTTACCTTAAGGATTAATTAATTAATG 449
Db 2874 TAAT-----TCAGCTGGAATAATCTAATTAAGGATTAATGATTAATGATTAATG 2918
OY 450 GACTTTAAAGATCCGCGGAGAGTATGACAAATTAATTAATTTTGGGATTAATGATTAATG 509

```

```

Db 2918 GACTTTAAATGATTAATTAATGAAAAACCAATCAGTATTTTGAATATTAACATTAAGCA 2978
OY 510 ATTATATGCTTATTTAGCAATTAATGAGTGTGTTTAACTATTTACTATTAATGATATATC 569
Db 2979 AGATATATCAGGATATTAATTAATGATGTTTGTGTTTAACTATTTACTATTAATTAAT- 3035
OY 570 TTCTGCTAATTTGTAATTAATGAGTACTTATGGAAGTGCAGAAATTAATGATGTTAG 629
Db 3036 TAATCTAAATTTTATTAATTAAGTACGCTTATGATTAATGATTAATTAATTAATGATTA 3095
OY 630 AGCTTTAGAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 689
Db 3096 AGAATTTATGTTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3155
OY 690 ATACGTTCTATTTGATTAATTTAGATATTTGCAACAGCATTAATTAATTAATTAATTA 749
Db 3156 ATTATTTGATGAATATTTTATTTAGATTTTAAATGCAATTAATTAATTAATTAATTA 3215
OY 750 AAAATTAACAGAGTTATTTATCTATTAACCTTTTAAAGACTGTCGGGAAACCTTT 809
Db 3216 AGAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3275
OY 810 ACGATATGATCAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 869
Db 3276 AATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3335
OY 870 GAAATATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 929
Db 3336 AGTGAAGATTTCACTCTAGGTGAAATATTAATTAATTAATTAATTAATTAATTAATTA 3395
OY 930 GAATATATTTATTAAGAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 987
Db 3396 TTATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3455
OY 988 -CCTAATTAATGAATGATTTCTTTTAAATCAGGATTTTAAATTAATTAATTAATTA 1046
Db 3456 TTCTCATCTTAATTAATGATATATGATTAATTAATTAATTAATTAATTAATTAATTA 3515
OY 1047 ATATTAACAAATTAATGAGACATTTAGGTATCCGAAAGATGAATG--CCTTAATTA 1103
Db 3516 ACTTCAACATGAAGAGTGGAGATATATGCTTAATTAATTAATTAATTAATTAATTAAT 3575
OY 1104 TCTTATGATGATTTCTAAGATAGTTATTAATTAATTAATTAATTAATTAATTAATTA 1163
Db 3576 ATTGTTTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3635
OY 1164 GGAAGCAGTAA--ATGCGTATTTTAAACCTATTTCTGACCACTTAATTAATTAATGA 1220
Db 3636 ATATGATGAACGCCATCATATATGATGTCAGTTTAAATTAATTAATTAATTAATTA 3695
OY 1221 TGATAAAATGATCTTTTGAAGTACTAGTACGCTTAAATTAATTAATTAATTAATTA 1280
Db 3696 TGATGATATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3755
OY 1281 AATATGAGTATATTAATTTGCAAGCAACTGTACTTAATTAATTAATTAATTAATTA 1334
Db 3756 GTATTAAGATTTTATTTGATTAAGTAAATGATTAATTAATTAATTAATTAATTAAT 3815
OY 1335 -----AATTTAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1382
Db 3816 TAATGAATTTGGGATGTAATTTGGCAGTTTATTCCTAAAGATGAAGGTGAGC 3869

```

RESULT 4

```

US-10-051-952-1
; Sequence 1, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17

```

Oy 626 TAGGACCTATTAGAGAGATTAATAATTAACATTAAACCTAGATAGCTAATAATAATA 68
 |||| |||| | ||||||||| |||| |||| ||||
 Db 3488 TAGCTAATATTTCATGCTAGTAATAATAATAATGTTTAAATAGATGCTGTTGAGACATTACAC 35

QY	266	TTCAATATATGATATGTTAAATTAATTTTACCCTAGCTTGGTGGTGGGTTCCCTAAAG	325
Db	1366	TTGTATATATAGTATGCTATGAAAATTTTGTACTTGGTTTGGATTAAGAAATTCCTAAGT	1425
QY	326	TATCTGCTAGCTATTTAGAACAAATATGCGCAAAATGAGTATTCATATATTAAGTCTGTAGA	385
Db	1426	ATTTTAAACAGT-----ATTAAGTCTAAATATATGATATACAAATATATAATTTGTAGT	1476
QY	386	AAAAACATATGCTATCAATAGACTCTGGTGGAGTGTATCATTAAAGTATATAACTTTAA	445
Db	1477	AAATATAT-----TCAGGAATGGAAGATGATCATTAAATTTATGGTAAATTA	1521
QY	446	TATGACTTTAAAGATATCCGCGGAGAACTTTGACAAATTAACCTTTATAGGATTTTACTCG	505
Db	1522	TCGGACTTTACGAGGATCTACAGGAAATTAACAAGAGTATGTTTTTAATA-----	1573
QY	506	ATTAATTTATGCTTATTTAGCAAAATAATGGTTTTTATTAACATTACTAATGATAGAT	565
Db	1574	--CAGTCAATATGATTAATATATCAACAGATGGATTTTTTGTAACTATCATAATATAGAT	1631
QY	566	TATCTCTGCTATTTGTATATAAATGGAGTACTTATGGCAAGTGCAGAAATTAATCTGTT	625
Db	1632	TAAATTAACCTTAATAATTTTATATAAAGGAAGATTAATAGATCAAAAACCAATTTCAATT	1691
QY	626	TAGAGCTTATAGAGAGGATATAATATATAACATTAAGAATCTAGATGTAAATATATATA	685
Db	1692	TAGGTAAATTTCAATGCTAGTAAATAATATATATGTTTAAATTTAGATGCTTGAAGAGATAC	1751

Db	1752	ATAGATATATTTGGATAAATATTTTAACTCTTTTATAGAAGATTTAAATGAAAAGAA	1811
QY	746	TTGAAANAATTATACACAGTTATTTATCTATAACCTTTTAAAGACCTCTGGGGAACC	805
Db	1812	TCAAAGATTTATATGATATATCAATCAATCAAAATTCAGTATTTTAAAGACCTTTGGGGTGATT	1871
QY	806	CTTTACGATATATACACAGATATTTATTTATACACAGTACCTCTAGCTCTGTCGTAAGATGTC	865
Db	1872	ATTTACATATATGATTAACCATCTATCTATCTTAATTTATATGATCCAAATTAATATGTGC	1931
QY	866	AATTGAAAA-----TATACAGATTAATATGATTTGCAAAATGGCCATCGTATACTA	919
Db	1932	ATGTAATAAATAGTAGGTATTTAGAGGTTATATCTTAAAGGCGCTAGAGTAGGCTAA	1991
QY	920	ACGGAAATTTGATATATTTATTTAGAGGTTATATATAGSACTAAATTTATATAAAA	979
Db	1992	TGACTACAAACATTTATTTAAATTTAAAGTTTGTATATAGGGGACCAAAATTTATATAAAA	2051
QY	980	GATATACACCTTAATATAGTAATAGATTCCTTTTGTAAATCAGAGTATTTATTAATTAAT	1039
Db	2052	AAATATGCTTCTCGAAAT---AAAGATTAATATTTGTAGAAATATGATCGTATATATTA	2108
QY	1040	ATGATATCAATTAACCATTAATAGCGCATTTAGGT	1074
Db	2109	ATGTAGTAGTTAAAAAATTAAGATTAATAGCTTAGCT	2143

```

RESULT 7
US-09-350-756-9
: Sequence 9, Application US/09350756
: Patent No. US20020034521A1
: GENERAL INFORMATION
: APPLICANT: U.S. Army Medical Research Institute for Infectious
: Diseases
: APPLICANT: John S. Lee
: APPLICANT: Peter Pushko
: APPLICANT: Michael D. Parker
: APPLICANT: Jonathan F. Smith
: APPLICANT: Mark T. Dertzbaugh
: APPLICANT: Leonard Smith
: TITLE OF INVENTION: Botulinum Neurotoxin Vaccine
: FILE REFERENCE: 003/124/SAP R1ID 98-21
: CURRENT APPLICATION NUMBER: US/09/350 756

```



```

Db      853 TATCAGAAACCAAT--ATTTTCCACACCTAGATTATATACAGAGTATATTA 910
Qy      974 TAAAAA-----GATATACACCTATATATGAATAGATTCTTTGTTAAATCAGCTGATT 1027
Db      911 TAAGAAAAAATGATCTACAGATATATCTAATACAGATAATTTGTGTGAAAAAATGATC 970
Qy      1028 TTATTAATATATATGATCATAT 1050
Db      971 TGGCATATATTAATAGTAGTAGAT 993

RESULT 10
US-10-051-952-6
; Sequence 6, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-6

Query Match      14.88; Score 205.6; DB 12; Length 3753;
Best Local Similarity 51.28; Pred. No. 4.2e-28;
Matches 658; Conservative 0; Mismatches 584; Indels 42; Gaps 6;

Qy      16 CCATTCATTTCTTATCTTAAATCTGATGTTGGTTGATATGAAAGATATA 75
Db      2470 CCTTTAAAGCTTCTCTTATACAGATGATAAATTTTATTCATTTTAAATAGTTTC 2529
Qy      76 GATGTTATATTAAGAGATCAATTTTAAATTTAGATATTAATATGATATTTATATCA 135
Db      2530 TTTAAGAAATTAAGAGATGTTCTGTTTAAATATGAGATATAAAATGATAAATAGTA 2589
Qy      136 GATATATCTGGTTTAAATCTGTAATACATTCATCCAGATGTCATTTGGTCCCGA 195
Db      2590 GATATCTAGATATGATTTCAATATTAATTAATGAGATGTAATTAATATCC---A 2646
Qy      196 ATAAATGCAAGCAATACATTTAGTAACAATGAAATCTTCTGAGATATAGTCATAA 255
Db      2647 ACTAATAAATCAATTTGGAATATATATGATTAAGTAACTAGTAAATATATCTCAA 2706
Qy      256 GCTATGATATTTGATATATATGATATGTTTATATTTTACCGTAGCTTTGGTGAAG 315
Db      2707 AATGATTAATATATATATGATATATTAATTAATTTAGTATTTTGGGTAAAGA 2766
Qy      316 GTTCCATAAGTATCTAGTACATTTAGAACAAATATGCGACAAATAGCATTTCAATATT 375
Db      2767 ATTCCTAA-----CTATGATATATAGATAGTAATGTTAAATATGAATACACTATAA 2820
Qy      376 AGCTCTATGAAAAACATAGTCTATCAATAGAGATCTGTTGAGAGTATACCTTAAGGT 435
Db      2821 AATGTATAGAGGATTAATAT-----TCAGATGGAAGATATCTCTTAATATCA 2868
Qy      436 AATACCTAATATGAGATTAATAAGATCCGGGGAAGATGAGCAATTAATCTTAGAG 495
Db      2869 AATGAAATATTTGAGCATATGCAAGATTAATTCAGGATTAATCAAAAATTTAGCATTTAAC 2928
Qy      496 GATTATACCTGATATTTAATGCTTATTTAGCAAAATTAATGAGTTTATTAATCTATTAC 555
Db      2929 TATGTAACGCAATGATATTTCTGATTTATTAATTAATAGTGGATTTTGTGAATATAACT 2988
Qy      556 AATGATAGATTATCTTCTGCTAATTTGTATATAATGAGTACTTATGGAGAGTGCAGAA 615

```

```

Db      2989 AATGATAGATTAGAGATTTCTAACTTATATTAATGGAATTTAATAGATAAAAATCA 3048
Qy      616 ATTACTGTTTGGAGCTATTTAGAGAGATATAATTAATTAACATTAAGATAGATGT 675
Db      3049 ATTTTAAATTTAGTATATTCATGTTAGTACAAATATATTAATTTAAATAGTTAATGT 3108
Qy      676 AATTAATATATCAATACGTTTCTATTTGATTAATTTAGATATTTTGCACAAATTAAT 735
Db      3109 AGTTATACAA---GATATATGTTATAGATATTTTATTAATTTTGTGATAAGATTTGAT 3165
Qy      736 CCAAAAGAGATTGAAAAATTAATACAGATTTATTTATCTATACCTTTTAAAGACTTC 795
Db      3166 GAAACAGAAATTCAAACTTATATTAACAAATGAACTTAATCAATATTTTAAAGATTTT 3225
Qy      796 TGGGGAACCTTTAGCATATGATATGATAGATATTTATTTAATACAGACTTCAGTTCT 855
Db      3226 TGGGGAATTTATTTGCTTTATGACAAAGATATGATATTTAATTTAATGT-GTTAAACCAA 3284
Qy      856 AAGATGTTCAATTTGAAAAATATACAGATTAATGATTTGACAAATGCGCCATCGAT 915
Db      3285 TAACTTTATTAATAGAGAGAACAGATTTCTACTTTAAGCATTA-----AT 3327
Qy      916 ACTAACGAAAAATGATATATTTATAGAAAGTTATATATAGTACCTAAATTTATTA 975
Db      3328 AATATGAAGAGCACTATTTCTTTAGCTAAATAGATTAATATAGTGAATGAATAAATTA 3387
Qy      976 AAGATATTCACCTAATATGAAATGATTTCTTTGTTAAATCAGGATTTTATTTAA 1035
Db      3388 CAAGAGTTATATATGTAGTACTACGATATATCTTTGTTAAAGATATCAGTATAT 3447
Qy      1036 TTATATGATATATATACAAATTAATGACACATTTAGTTATTCGAAAGATGAATGCC 1095
Db      3448 ATTAATTTTGTACGACGACAAACCTACTTCCATTAATATATGATAGACATACCA 3507
Qy      1096 TTTAATATCTTGTATGATATTTCTAGAGTATATATAGCCCAAGATATCCCTTTAT 1155
Db      3508 AATTAAGAGAAAACAATTAATATATCATCTGTCATAGATTTAATCAAGTACTAT 3567
Qy      1156 AAAAAATGAGACAGTAATAATGCGATTTAAAAACCTATTCGTAACCTTAATTA 1215
Db      3568 ATGATTTCAAGTATGATATACATGATTTTAAATATATATATGAAATATATTTGGTTG 3627
Qy      1216 TATGATATTAATAATGATCTTTTGAAGTATAGTACCTTAATGTCATTAATAGCAAC 1275
Db      3628 TTAGGTTTCAAGGAGATGATGTTGTTGTTGTTATTTATACATATAGAGAT 3687
Qy      1276 GATCCAAATAGGATATATTAAT 1299
Db      3688 AATACAAACAGCAATGATTTT 3711

RESULT 11
US-10-051-952-8
; Sequence 8, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 3825
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-8

Query Match      13.08; Score 181.4; DB 12; Length 3825;

```

Best Local Similarity 51.7%; Pred. No. 7.8e-24;
Matches 520; Conservative 0; Mismatches 466; Indels 19; Gaps 4;

```

QY 16 CCAATTCATTTTCTTATTTCTAAATATCGAGTGTGGGATGATATGAGAGATATA 75
Db 2524 CCATTTGAACTTCTCTCATATCTATATATTAATTTCTATATATTTTATATGATTA 2583
QY 76 GATGTTATATTTAAAGAGTACAAATTTAAATTTAGATATTAATATGATATTAATCA 135
Db 2584 TATATAAAATTAAGATATGTTCTATTTTATGATATGCGATATGAAATATATTAATTA 2643
QY 136 GATATATCTGGTTTATTCATCTGTATTAATCATATCCGATGCTCAATTTGGCCCGGA 195
Db 2644 GATATCTCTGATATGTTCAATATATAGCAATTAATGGAACGTATATATTTATTC--A 2700
QY 196 ATAAATGGCAAGACATATCAATTTAGTAAACATGATCTTCTGAGTTTATGATCAATA 255
Db 2701 ACAATATGAAATCAATTTGATATATATATAGTAGCTTATGCAAGTTATATAGTCA 2760
QY 256 GCTATGATATGATATATATATGATATGTTTATATTTTACCCTTATGCTTTGGTTGAG 315
Db 2761 ATATATGATATATATATATATATATATATATATATATATATATATATATATATAT 2820
QY 316 GTTCTTAAAGATCTGATGATTTAGACAAATATGCAAAATGAGATTCATTAATTT 375
Db 2821 ATTCCTAAA-----CACTAGAACCTATGATATCAATATCGGAAATAGACTATA 2868
QY 376 AGCTCTTGAAGAAACATAGTCTATCAATAGATCTGTTGGAGTGTATCACTTAAAGT 435
Db 2869 ATAAATTTGATGGGATATATATATGCGGATGGAATATACATAGAGCTTTAGAGT 2928
QY 436 AATTAATTAATGATGATTAAGATATCCCGGAGAGATAGTACAAATACCTTTAGG 495
Db 2929 TGCTCAATTAATTTGGACTTTCACAGATATCTCTGAAATTAAGAAATTTATTTTATAGG 2988
QY 496 GATTTACCTGATTAATTAATGCTTATTTAGCAATTAATGCTTTTATATCTTTTCT 555
Db 2989 TATCAAGACCTTATAGATATCTAATTTATATATATATATATATATATATATATATAT 3048
QY 556 AATGATGATATATCTTCTGCTAATTTGATATTAATTAATGAGTATGGAAGTCAGAA 615
Db 3049 AATTAATTAATGATGATGATCTAGATTTAGATCAATGGAATTTAATAGTTGAATAATCA 3108
QY 616 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
Db 3109 AATTCGATTTAGTGTATTTCAATTTAGTATATATATATATATATATATATATATATAT 3168
QY 676 AATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
Db 3169 GATGATGAA--CGTATGTTGATATAGATATTTTAAAGTTTAAATAGGATTTAGAT 3225
QY 736 CCAAAAGAGATGAAATTAATATACAGATTTATTTATCTATACCTTTTAAAGACTTC 795
Db 3226 AAAACGAAATTAAGACTTTATATATAGATGAGCAGATCCAAATATCTTAAAAACATAT 3285
QY 796 TGGGGAACCTTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
Db 3286 TGGGGAATTTATTTGATATATATATATATATATATATATATATATATATATATATAT 3345
QY 856 AAGATGTTCAATGAAATTAATATACAGATTTATATGAT--TTGACAAATGGCCATGTA 914
Db 3346 AAGTATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3405
QY 915 TACTATGGAATTAATATATATATATATATATATATATATATATATATATATATATAT 974
Db 3406 GGTCTGTTTATTTGATATATATATATATATATATATATATATATATATATATATATAT 3465
QY 975 AAAAATATATACATTAATATATATATATATATATATATATATATATATATATATATAT 1019
Db 3466 GGTCTATATATATATATATATATATATATATATATATATATATATATATATATATAT 3510

```

RESULT 12

US-10-051-952-5
; Sequence 5, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3831
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-5

Query Match 11.2%; Score 155.8; DB 12; Length 3831;
Best Local Similarity 51.9%; Pred. No. 2.5e-19;
Matches 451; Conservative 0; Mismatches 382; Indels 36; Gaps 3;

```

QY 9 TTCAACACCAATTCATTTCTTATTTCTAAATATCTGATGTTGGTGAATGAAGA 68
Db 2514 TACAATGCCCTTTAATTTTTCATATATATATATATATATATATATATATATATAT 2573
QY 69 AGATATATATGTTATATATATATATATATATATATATATATATATATATATATATAT 128
Db 2574 TGAATATTTCAATATATATATATATATATATATATATATATATATATATATATAT 2633
QY 129 TATATGATATATATATGTTGTTTATTTATTCATCTGATTAATCAATATCAATGCTCA 188
Db 2634 TTTAGTGATATATATATATATATATATATATATATATATATATATATATATATAT 2688
QY 189 GCCCGAATTAATGCAAGCAATACATTTAGTAAACATGAAATCTTGAAGTTATAGT 248
Db 2689 -CTTAATGATATATATATATATATATATATATATATATATATATATATATATAT 2747
QY 249 GCATTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308
Db 2748 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2807
QY 309 GTTGAAGGTTCTTAAGATATCTGCTAGTATTTAGAACAAATATGCAAAATGATGATTC 368
Db 2808 GATTAAGATATATATATATATATATATATATATATATATATATATATATATATAT 2852
QY 369 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 428
Db 2853 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2897
QY 429 TAAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 488
Db 2898 TAGGATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2957
QY 489 TTTTGAAGATTTTACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548
Db 2958 TTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3017
QY 549 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 608
Db 3018 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3077
QY 609 TGCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 668
Db 3078 TCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3137
QY 669 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 728
Db 3138 TGAGATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3197
QY 729 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 788

```

Db 3198 ATTAACTAATGAAGATATTATATGTATATGAGCAAAATATTAAGAATGTTATTA 3257
Qy 789 AGACTTCTGGGGAACCCCTTACGATATGATACAGATATTTATTAACCAGTACCTTC 848
Db 3258 AGATTTATGGGGAATCCCTTGAAGTTTGATACAGATATTTATTAATGATATTA 3317
Qy 849 TAGCTTAAAGATGTCATTTGAAAAATA 877
Db 3318 TATAGATAGCTATATTGACCTGAAAGTA 3346

RESULT 13

US-09-350-756-8
; Sequence 8, Application US/09350756
; Patent No. US2002004521A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research Institute for Infectious Diseases
; APPLICANT: John S. Lee
; APPLICANT: Peter Pushko
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Mark T. Dertzbaugh
; APPLICANT: Leonard Smith
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine
; FILE REFERENCE: 003/124/SAP RIID 98-21
; CURRENT APPLICATION NUMBER: US/09/350,756
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 8
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; FEATURE:
US-09-350-756-8

Query Match 10.4%; Score 145.2; DB 10; Length 1987;
Best Local Similarity 55.5%; Pred. No. 1.7e-17;
Matches 386; Conservative 0; Mismatches 273; Indels 37; Gaps 4;

Qy 86 TAAAAAGAGTACATTTTAAATTTAGATATTAATTAATGATTTATTCAGATATATCTG 145
Db 1319 TTATTAATAGCTTATATGATTTAGATTAAGAAATCAATTTAATGACTTATCTA 1378
Qy 146 GGTATTAATCTATCTATTAATGATATCCAGATGCTCATTTGTCGCCGGAATAATGCA 205
Db 1379 GGATGATCAAAAATAAATTTGGTAGTAAGTAATTTTGATCC--AATGATTAATA 1435
Qy 206 AAGCATATACATTTAGTAAACAAATGATCTCTGAAGTTATAGTCATTAAGCTATGATA 265
Db 1436 ATCAATTTCAATTTTATTAATTTAGAAAGTAAATTTAGAGTAAATTTTAAAAATGCTA 1495
Qy 266 TTGAATATATGATATGTTAAATTTTACCGTTAGCTTTTGTTGAGGTCCTTAAG 325
Db 1496 TTGAATATATAGATGATGAAATTTTACTAGCTTTTGGATGAAGATTCCTAAGT 1555
Qy 326 TATCTGCTAGCATTTAGAACAAATATGCGCAANTGATGATCAATATATGCTCTATGA 385
Db 1556 ATTTTAAACAGT-----ATAGCTCAATTAATGATTAACAAATTAATTTGATGG 1606
Qy 386 AAAAACAATAGTCTATCAATAGATCTGTTGGAGTGTATCACTTAAGGTAATAACTTAA 445
Db 1607 AAAATTAAT-----TCAGATGGAAGATATCACTTAATATATGCTGAATTA 1651
Qy 446 TATGACTTTTAAAGATTTCCGGGAGAGATTTAGAACAAATTAATTTTATGAGATTTACTG 505
Db 1652 TCTGACTTTTACAGGATACCTGAGAAATTAACAAAGATGTTTAAATATACAGTCAA 1711
Qy 506 ATTAATATATGCTTATTTAGCAATTAATGCTTTTATTAATGATTTATGATGATAGAT 565
Db 1712 TGATTAATATATACATTAATTAATTAACAGATGATTTTGTAGACTTATCACTAATATAGAT 1771

Qy 566 TATCTCTGCTAAT-----TGATATAAATGAGTACTTATGGAAGTGCAGAA 615
Db 1772 TAATTAATCTTAATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1831
Qy 616 ATTACTGTTTAAAGACATTTAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 675
Db 1832 ATTCAATATTTAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1891
Qy 676 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 735
Db 1892 AGAATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1951
Qy 736 CCAAAAGATTTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771
Db 1952 GAAAAAGAAATCAAAAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1987

RESULT 14

US-10-051-952-4
; Sequence 4, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3876
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-4

Query Match 8.2%; Score 113.8; DB 12; Length 3876;
Best Local Similarity 49.7%; Pred. No. 6.5e-12;
Matches 481; Conservative 0; Mismatches 422; Indels 64; Gaps 5;

Qy 9 TTCAACCAATTCATTTCTTTATTTCTAAATCTGATGTTGGTGTAAATGAAGA 68
Db 2526 TACATATCCCTTAATTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 2585
Qy 69 AGATATGATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128
Db 2586 TGAATATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2645
Qy 129 TATATCAGATATATCTGGTTAAATCTGTAATTAATTAATTAATTAATTAATTAATTA 188
Db 2646 TTATGATGATCAATCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2705
Qy 189 GCCCGGAA-----AATGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTA 245
Db 2706 TCCATATTTCCATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2765
Qy 246 AGTCATTAAGCTATGATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 305
Db 2766 AGTAACCCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2825
Qy 306 TTGTTGAGGTTCCCTTAAGATATCTGCTAGCATTTAGAACAAATTAATTAATTAATTA 365
Db 2826 TTGATTTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2857
Qy 366 TTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 425
Db 2858 -----TACCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2912
Qy 426 ACTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 485
Db 2913 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2972

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:57:42 ; Search time 1951 seconds

(without alignments)
11555.160 Million cell updates/sec

Title: US-09-816-467-3

Perfect score: 1392

Sequence: 1 atggttttttcaaccacaaat.....aggagatgacaaatgatataa 1392

Scoring table: IDENTIFY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estcpl:*
7: em_estro:*
8: em_esthc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estom:*
16: gb_gss:*
17: em_gss_hum:*
18: em_gss_pln:*
19: em_gss_vtl:*
20: em_gss_fun:*
21: em_gss_mus:*
22: em_gss_other:*
23: em_gss_pro:*
24: em_gss_trod:*
25: em_gss_rod:*
26: em_gss_rod:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.2	7.3	1101	17	CNS00DEVL
2	92.6	6.7	1101	17	AL069706 Drosophila
3	89.2	6.4	1190	17	CNS00396
4	89	6.4	1201	17	CNS020N7
5	88.4	6.4	1190	17	CNS0167M
6	87.4	6.3	1101	17	CNS00EVL
					AL069706 Drosophila

7	87	6.2	1101	17	CNS003BD	AL064091 Drosophila
8	83.6	6.0	1101	17	CNS00E07	AL069440 Drosophila
9	80.8	5.8	1001	17	CNS0155H	AL050523 Drosophila
10	80.2	5.8	876	17	CNS009G1	AL053529 Drosophila
11	79.4	5.7	829	17	CNS03L0B	AL049932 Tetradon
12	79.4	5.7	890	17	AQ026918	AQ026918 CTF-HSP-2
13	78.6	5.6	1101	17	CNS00EPO	AL069493 Drosophila
14	78.2	5.6	1101	17	CNS0039G	AL069440 Drosophila
15	78.2	5.6	1101	17	CNS00E07	AL040900 T3 end of
16	77.8	5.6	1169	17	CNS06KHO	AL071206 Drosophila
17	77.6	5.6	1101	17	CNS00EFG	AL063632 Drosophila
18	77.4	5.6	1027	17	CNS02750	AL071233 Tetradon
19	77.4	5.5	1101	17	CNS000B8	AL069440 Drosophila
20	77.4	5.5	1101	17	CNS000B8	AL069440 Drosophila
21	76.4	5.5	820	17	B11728	AL0425139 clone BA0
22	76.2	5.5	897	10	BB620968	BB620968
23	75.6	5.5	1084	17	CNS071NH	AL077628 Drosophila
24	75	5.4	1101	17	CNS00KAE	AL077628 Drosophila
25	74.8	5.4	893	17	BF274559	AL077628 Drosophila
26	74.8	5.4	1092	12	BF274559	AL077628 Drosophila
27	74.6	5.4	1101	17	CNS002K7	AL077628 Drosophila
28	74.4	5.3	581	17	CNS034DK	AL077628 Drosophila
29	74.2	5.3	1152	12	BC309087	AL077628 Drosophila
30	74	5.3	1187	17	B11102	AL077628 Drosophila
31	73.2	5.3	928	17	CNS00DKY	AL077628 Drosophila
32	73.2	5.3	1161	17	CNS073Y8	AL077628 Drosophila
33	73	5.2	923	17	AG058591	AL077628 Drosophila
34	73	5.2	1225	17	CNS0161D	AL077628 Drosophila
35	72.6	5.2	1101	17	CNS001FB	AL077628 Drosophila
36	72.4	5.2	734	17	CNS010MP	AL077628 Drosophila
37	72.4	5.2	961	17	CNS008HT	AL077628 Drosophila
38	72.2	5.2	994	17	CNS04N0T	AL077628 Drosophila
39	72.2	5.2	1006	17	CNS00813	AL077628 Drosophila
40	72	5.2	945	17	CNS04DOK	AL077628 Drosophila
41	71.8	5.2	976	17	CNS04E5M	AL077628 Drosophila
42	71.8	5.2	1101	17	CNS00EFG	AL077628 Drosophila
43	71.6	5.1	895	17	CNS00EFG	AL077628 Drosophila
44	71.4	5.1	867	17	CNS075BG	AL077628 Drosophila
45	71.4	5.1	923	17	AG058591	AL077628 Drosophila

ALIGNMENTS

RESULT 1
CNS00EVL
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL069706
GI:4949849

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial

[illegible]

RESULT 3	CNS020N7	1190 bp	DNA	linear	GSS 14-MAY-2000
LOCUS	CNS020N7				
DEFINITION	Tetraodon nigroviridis genome survey sequence PUC-ori end of clone sequence of library G from Tetraodon nigroviridis, genomic survey				
ACCESSION	AL206908				
VERSION	AL206908.1				GI:7865727
KEYWORDS	GSS; genome survey sequence.				
SOURCE	Tetraodon nigroviridis.				
ORGANISM	Tetraodon nigroviridis				

REFERENCE	TITLE	JOURNAL	REFERENCE
1	Roest-Crolius/H., Jallion/O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottilier,P., Queclier,F., Saurin,A. and Weissenbach,J., Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence	Unpublished	Roest-Crolius/H., Jallion/O., Dasilva,C., Fizames,C., Fisher,C., Bernot,A., Wincker,P., Brottilier,P., Queclier,F., Saurin,A. and Weissenbach,J., Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
2	(bases 1 to 1190)		

TITLE	Characterization and repeat analysis of the compact genome of the freshwater putterfish <i>Tetraodon nigroviridis</i>
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1190)
AUTHORS	Genoscope
TITLE	Direct Submission
JOURNAL	Submitted (12-Apr-2000)
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the <i>Tetraodon nigroviridis</i> genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .

FEATURES	
source	Location/Qualifiers
	1..1190
	/organism="Tetradodon nigroviridis"
	/db_xref="taxon:998883"
	/clone="153P04"
	/clone_11b="G"
	/note="genoscope sequence ID : C0AG155DH02SP1-end : PUC-Or1"
BRASE COUNT	373 a 229 c 141 g 312 t 135 others
ORIGIN	

Query Match Similarity 6.4%; Score 89.2; DB 17; Length 1190;
Best Local Similarity 42.2%; Pred. No. 5.7e-07;
Matches 188; Conservative 56; Mismatches 200; Indels 2; Gaps 2;

QY	684	TATACGATCGCTTCGATTGATAAATTTAGCATATTTGGCAAGCATTAAATCCAAAAGA	743
Dd	643	AAATTTAAMAAAAATDAATDATTAMTNWMAAATAAANAAAAAATATTATATAMNATAMNWAA	702
QY	744	GATGTAAAAATTTATPACAAGT-TATTATTCTATAACTTTTAAAGAGACTTCGGGGA	802
Dd	703	TATATATTATATATWATVANNPATPATATATATANAATNTATTATTTTTTTTTNNAT	762
QY	803	ACCCTTAGCATATGTATACAGAATATTATTTAAATPACAGTAGCTTGCTCTAAGAG	862
Dd	763	AATTTNTTMAWWMTATTAATWTATAATANTAWMTAWTNTTWTTTATATATATAMWWW	822
QY	863	TTCAATTGAAAAATATPAACAGATTAATATGATTGTGCAGATGGCCATGCTATACTAACG	922
Dd	823	TATWTTTTTTTATATATWWTWTATTTWTATIMWAATTTTAAAAATATNTTATATPAAAAWATA	882
QY	923	GAAATTTGAATATATATTTATAGAGGTTATATATGAGCTATAAATTTAT-TTPAAAAAGA	981
Dd	883	TAAATATATNTATATANNMWTANNAATATAMAATATATATATAMMATITATATATATATATA	942
QY	982	TPTACACCTTAATATAATGAATAGATICTTTTGTGTTAAATCAGTGATTTTATTAATATTAT	1041
Dd	943	TAAATATATATWCAYATINATWTTATATTTAAAMATANATATATATVATPMBAACAAATATATAT	1002
QY	1042	GTATCATATATAATATATGACACAT	1067
Dd	1003	CACAAACAAACATACAAAMMNATAT	1028

RESULT 4	
LOCUS	CNS0167M/c
DEFINITION	CNS0167M 1201 bp DNA linear GSS 26-JUL-1999
ACCESSION	Drosophila melanogaster genome survey sequence T7 end of BAC
VERSION	BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit
KEYWORDS	fly), genomic survey sequence.
SOURCE	AL106396
ORGANISM	GI:5621701
REFERENCE	GSS.
AUTHORS	Drosophila melanogaster.
TITLE	Drosophila melanogaster
JOURNAL	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1201) Genoscope.
COMMENT	Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

```

FEATURES
Source      Location/Qualifiers
1.1201
/organism="Drosophila melanogaster"
/db_xref="taxon:7237"
/clone="BACN15M24"
/cione_1b="DrosBAC"
/plasmid="pbeloBAC11"
/note="end : T7"

BASE COUNT      323 a      87 c      79 g      551 t      161 others

ORIGIN

```

Query Match	6.48;	Score 89;	DB 17;	Length 1201;
Best Local Similarity	37.0%;	Pred. No. 6.2e-07;		
Matches 273;	Conservative 90;	Mismatches 361;	Indels 14;	Gaps 1;

```

      1187  WATATATATAATATAAATAATATAMWMAATATATAMNMAAANTATATAAATAAATAAATAATAM 1128
QY      696  TCTCTATTTGTAATTTAGAGATTTTTCGAAGCATTAATCCAAAGAGATTGAAAAATT 755
Db      1127  WATATATAAATAAATAAATAATATAATATAATATAAATAAATAAATAAATAAATAAATAA 1068
QY      756  ATACACAGTTATTTATCTATACCTTTTAAAGACCTTGGGGAACCTTTACGATA 815
Db      1067  TTTANATATTTTWTWNTATATAATATTTTATTTTATTTTATATAAATAAATAAATAAATAA 1008
QY      816  TGAATACAGATATTTATTTATACCACTAGCTCTACTTCAAGATGTCATGCAATGAAAA 875
Db      1007  AATTTTAAAAATAAATAATATATAAATAATTTTAAAAATTTTATATTTTATTTTAAAAA 948
QY      876  TATACAGATTTATATGTTATTTGCAAAATGGCCATCTATACGAAATTAATGATAT 935
Db      947  AAAAATAATATWMAAATTTTATTTATTTATATAAATAAATAAATAAATAAATAAATAAATAA 888
QY      936  ATATTTATAGAGTTATATATATGACTAAATTTATTTATATAAAGATATACCTATAATA 995
Db      887  AAAAATTTTAAAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 828
QY      996  TGAATAGATCTTTGTTGTAATCAGCTGATTTTATTAATTTATATGATATATACAA 1055
Db      827  TAAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 768
QY      1056  TAAATGACACATTTAGAGTTATCGGAAGATGGAATGCTTTATATATCTTGTAGAT 1115
Db      767  AATATWMTAAATATATATTAATTAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAA 708
QY      1116  TCTAAGAGTGTATATATGCCCCAGCTATCCCTTATATAAATAAATAAATAAATAAATAAATAA 1175
Db      707  AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 662
QY      1176  ATTCGCTGATTTAAAACTATCTGTACAATTAATTTATGATGATATAAATAAATAAATAAATAA 1235
Db      661  MAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 602
QY      1236  TTTAGAGCTAGTATGATCCCTATATGTCGAATAGCAGACATCCAAATAGGATATAT 1295
Db      601  TAAKNDKKMAADAKDANAKSANGBKMAATVACACTGGCGBAATAATGTAATTTGGC 542
QY      1296  AATTCAGACACTGTACTTATCATTTTAAAGATATAATTTTAAAGATGATGATGTA 1355
Db      541  AAASAAARRRAADDDGKKAARAARAKTCAAAAGAAAGCCTTAAGTTTTCATTTTC 482
QY      1356  CTTTGTACTACAGATGA 1373
Db      481  AATTTCCCTTTTCAGTTGA 464

RESULT 5
CNS020N7/c 1190 bp DNA linear GSS 14-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 13p04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL206908
VERSION AL206908.1 GI:7865727
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1190)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Winkler,P., Brotlier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1190)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1190)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-Apr-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..1190
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="153p04"
/clone_lib="G"
/notes="Genoscope sequence ID : C0AG153D02SP1-end :
PUC-ori"
BASE COUNT 373 a 229 c 141 g 312 t 135 others
ORIGIN
Query Match 6.4%; Score 88.4; DB 17; Length 1190;
Best Local Similarity 39.7%; Pred. No. 8e-07;
Matches 238; Conservative 65; Mismatches 294; Indels 2; Gaps 2;
QY 468 GGGAGAAGTTAGCAAAATTAACCTTTAGGATTTACCGATTAATTTAATGCTTATAGC 527
Db 1040 GAKAGCGATRTATATATWTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 981
QY 528 AATTAATGGCTTTTAACTATTTACTAATGATAGATTAATCTCTGCTAATTTGATAT 587
Db 980 TATTTATTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 921
QY 588 AATGCGACTTATGGAAGTGCAGAAATTTACTGTTTACGACTTATACAGAGATA 647
Db 920 ATATATATWATATATTTATATATATATATATATATATATATATATATATATATATATAT 861
QY 648 TAAATTAACATTAACAGTATGATGATATATATATATATATATATATATATATATATATAT 707
Db 860 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 802
QY 708 ATTAGATATTTTGAAGCATTAATCAAAAGATGGAATTAATTAATTAATTAATTAATTAAT 767
Db 801 WTATWATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 742
QY 768 TTTATCTATTAACCTTTTAAAGACTCTGGGGAACCTTACGATGATAGAGATA 827
Db 741 ATATATTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 682
QY 828 TTATTTATACAGTATGCTTACGATTTAAAGATGTCATGTAATTAATTAATTAATTAATTAAT 887
Db 681 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 622
QY 888 TATGATTTGCAAAATGGCCATCTATACGGAATTAATTAATTAATTAATTAATTAATTAAT 947
Db 621 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 563
QY 948 GTTATATATGACGTAATTTATATATAAAGATATACCACTAATTAATTAATTAATTAATTAAT 1007
Db 562 TTAATAAATAATTTTAAATATATATATATATATATATATATATATATATATATATATATAT 503
QY 1008 TTTTGTATATACGCTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1066
Db 502 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 444

```

RESULT 6

QY	571	TCTGTAATTTGTATAAATGAAGACTTACGGGAAATTCGTGGTAGGA	630
Db	607	TT	666
QY	631	GCTATTAGAGGAGTAATATAATTAACATTAAAACTAGATAGATGTAATAATAATCA	690
Db	667	TTTANNAAMWATTAATAAAAAAAAAAAAAAAAAAAAATATAAAAAAAAAATTAAMWTFA	726
QY	691	TACGTTCCATTTGATTAATTTTGAGATATTTCGCAAGCATTTAAATCCAAGAATGGA	750
Db	727	ATATATTMTTWTTTAAATAATATATTAATAAAAAAAAAAWATMTWTTTATTTATTTAATAAATAATMTAT	786
QY	751	AAATTTATCACAGTATTTATTTATCTAACCTTTTATAGAGACTTCGGGAAACCCTTA	810
Db	787	WATTTATATATNMNNAFTMTWTTTWTAAATTTWAATWTTAAATW-TMATTAAAMATTNTA	845
QY	811	CGATTTGATACGAATATTATTATTAATPACCAGTAGCTTCTAGTTCTPAAGATGTCATTTG	870
Db	846	WAAATTAATAAAAAAAAAAATAATTAATAATATATWATATWATMTWTTTAAATAATTAATMAATTA	905
QY	871	AAAAATATACAGATTTATGTATTTGACAATGGGCCCTCGTATACATACAGGAATAATG	930
Db	906	AAMAAAAAAATTTAAWATTAATAAATTAATTAATWTTTWTMYTMAAAWMTWATTAAAMAAMA	965
QY	931	AATATATTTTATAGAGGTTATATAATGAGCTAAATTTATTTATAAAAAGATATACACCT	990
Db	966	TATTTTGTATWATTAATTAATAAAMAAAAAAATAAATAAATAATWAAATNMWAMWVAAAM	1025
QY	991	AATATGAAATAGATCTTTTGTTAATCAGSGTATTTATTAATTAATATATATCATAT	1050
Db	1026	AYACMAAAAMATTTATWYATAMMAAAAAAAMHTTNATTTATTAWAAAMWAAHMAWCWTAT	1085

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CNS000E07	1101 bp	DNA	linear	GSS	04-JUN-1999							
<p>Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR29P01 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.</p> <p>AL069440</p> <p>AL069440.1 GI:4949583</p> <p>GSS.</p> <p>Drosophila melanogaster.</p> <p>Drosophila melanogaster.</p> <p>Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila.</p> <p>1 (bases 1 to 1101)</p> <p>Genoscope.</p> <p>Direct Submission</p> <p>Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 101 91006 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr</p> <p>- Web : www.genoscope.cns.fr)</p> <p>Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammosser in Pieter de Jong's Laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp. the same strain used for the BDGP's 1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.</p>												

FEATURES	location/Qualifiers
source	1..1101
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone="BACR29p01"

		/clone_lib="RPCT-98"			
		/note="end : TET3"			
BASE COUNT	366 a	66 c	104 g	351 t	214 others
ORIGIN					
Query Match,	6.0%;	Score 83.6;	DB 17;	Length 1101;	
Best Local Similarity	35.3%;	Pred. No. 6.2e-06;			
Matches 189;	Conservative 85;	Mismatches 261;	Indels 0;	Gaps 0;	

OY	3529	TTATGGGAGCTCCAAATAATACGGTTTGGAGCGTATTAGACAGGATATATATATTAACAT	658
Db	494	TAATTAATAAATTTAAAAAATAAATAATTAATAAATCAATWTAAATTAATTTTAAATTA	553
OY	659	TAAACCTGATAGATGTATATAATATATATCAATACGTTTCATTCATTAATTTAGCATAT	718
Db	554	WMAATKTTTTTATTATWMAATATATAAAMAATTTATTTTATTTATATTAATTAAGAA	613
OY	719	TTTGGCAAGCATTTAATCCAAAGAGATTGAAAATTTATACACAAGTTATATCATATA	778
Db	614	AAAAAAMTATAMHTHMAATTTATATMTWMAATTTAAATAMATTTTAAATTAAMT	673
OY	779	CCTTTTAAAGAGACTTCGGGAAACCTTTACGATATGATATACAGATATTTTAAATAC	838
Db	674	WTTTAAMWTTTAAATATATAAARWMAATTTMAAATAAAMWTTTMAATTAAMAAT	733
OY	839	CAGTAGCCTTCTAGTCTAAACATGTCATTAATGAATAAATAATACAGATATATGTTTGA	898
Db	734	AAATTTTTTTATTATTAAMWMAATTTAAATTAATMTATWMAAMWTTTAAMAAAAAATNTFA	793
OY	899	CAATGCCCCACVCGTACTATACGGAATAATGGAATATATATTATAGAGGTTATATAATG	958
Db	794	TAAATATWMAATATATWATATAMWTTTAAATATWATRTAAATTAATATWTTTWTWTA	853
OY	959	GACTTAATAATTTATATATAAAGATATACACTAATTAATGAATAGATCTTTGGTTAAT	1018
Db	854	TTTAAATTTTTTTATTAATWMAAAMWGTATATWAAATTTAMWTTTAAAMAATAATGTTTATC	913
OY	1019	CAGGTATTTTATTAATTTATATATGATCATATTAACAATAATAGACACACTTGTAGTTAT	1078
Db	914	TWMTCTATWTTTATTTTAAATTTGATATATAAAMAAMWATATGATATMMAAATRTWWTG	973
OY	1079	CGAAGATGGAAATGCCCTTAATAATCTTGATAGATCTTAAGACTACGTTATATA	1133
Db	974	TATTTKAAATATWGAATATTTAAATAAAAAAATAATTAAMAAMRTKGTGRTWTKA	1028

RESULT 9	CNS0155H	1001 bp	DNA	linear	GSS 26-jul-1999
CNS0155H					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence s96 end of BAC BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL105023				
VERSION	AL105023.1	GI:5617037			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1001)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :				
JOURNAL					

COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CERH (centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector

[illegible]

REFERENCE	TITLE	JOURNAL	REFERENCE	TITLE	JOURNAL
1 (bases 1 to 829)			2 (bases 1 to 829)		
Roest-Crollius,H., Jallou,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J.			Roest-Crollius,H., Jallou,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.		
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence			Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis		
Unpublished			Unpublished		
3 (bases 1 to 829)					
Genoscope.					
Direct Submission					
Submitted (12-APR-2000)					
This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .					

						/organism="Tetradodon nigroviridis" /db_xref="taxon:99883" /clone_"036F11" /clone_1Id="G- /note_"Genoscope sequence ID : CDBG036CC06SP1-end : PUC-Orl"						
BASE COUNT	594	a	0	c	2	g	232	t	1	others		
ORIGIN												
Query Match	5.7%	Score	79.4;	DB	17;	Length	829;					
Best Local Similarity	46.1%;	Pred. No.	3.9e-05;									
Matches	3337;	Conservative	0;	Mismatch	391;	Indels	3;	Gaps				
Oy	499	TTAACTGGATTAATTAAATGCCTATTAGCAATAAATACGGCTTTATATACATATTCATAT	558									
Dd	95	TAAATTAATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	154									
Oy	559	GATAGATTATATCTTCGCAAAATTTGTATTAATGAAGACTATATCGGAAGTGACAGAAAT	618									
Dd	155	AATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	214									
Oy	619	ACTGCTTTAGGACGCTATTAGAGAGCATTAATATTAACATTAAAACTAGATGATTAAT	678									
Dd	215	AATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	274									
Oy	679	AATATATATCAATACGCTTCTATGATTAATTTAGCATATTTTGCAGAACATTAATATCCA	738									
Dd	275	AATA--AATAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	332									
Oy	739	AAAGAGATTGAAAAATTTACACAAGTTATTATCTATAACCTTTTAAAGACCTTCGG	798									
Dd	333	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	391									
Oy	799	GGAACCCCTTTCGATATGATCAGACAAATATTTTAATATACCAGTAGCTTCTAGTTCTAAA	858									
Dd	392	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	451									
Oy	859	GATGTCATTAAGAAAATATATAAGATTATATGATTTGACCAAAATGCCGACATCGTACT	918									
Dd	452	ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	511									
Oy	919	AACGGAATTTGAATATATATATTATAGAAGGTTATATAATGACATAAAATTTATTAATA	978									
Dd	512	ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	571									
Oy	979	AGATATACCCATTAATTAATGAATAGATTCCTTTGTAAATCCAGTGATTTTATTAATTA	1038									
Dd	572	ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	631									

Qy	1039	TATGATCATTAATTAACAAATTAATGACGACATTTGAGGTTTCCGAAGAATGAATCCCTTT	1098
Db	632	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	691
Qy	1099	TATAATCTTGATAGAAATTCCTAGAGTAGGTTATAATGCGCCAGCTATCCCTCTTTATAAA	1158
Db	692	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	751
Qy	1159	AAATATGGAACGTAATAATTCGCTGATTTTAAAAACCTATCTGTGCAACTTAATTAATAT	1218
Db	752	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	811
Qy	1219	GATGATTAATAA 1229	
Db	812	ATTAATTAATA 822	

RESULT 12	890 bp	DNA	linear	GSS 30-JUN-1998
LOCUS				
DEFINITION	AOQ26918	CIT-HSP-2322B22.TF	CIT-HSP Homo sapiens	genomic clone 2322B22, DNA
ACCESSION	AOQ26918			
VERSION	AOQ26918			
KEYWORDS	GSS			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 890)			
	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,			
	Golden,K., Heriy,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,			
	Simon,H. and Venter,J.C.			
	Use of a random BAC End Sequence Database for Sequence-Ready Map			
	Building (1998)			
	Unpublished (1998)			
JOURNAL	Contact: Mark Adams			
COMMENT	Department of Eukaryotic Genomics			
	9712 Medical Center Dr., Rockville, MD 20850, USA			
	Tel: 301 838 0200			
	Fax: 301 838 0208			
	Email: mdadams@tigr.org			
	Closes are available from Research Genetics (info@resgen.com). BAC			
	end search page:			
	http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.			
	Seq primer: M13-21			
	Class: BAC ends.			

FEATURES	source	Location/Qualifiers
		1..890
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="232B22"
		/clone_id="CIT-HSP"
		/sex="Male"
		/cell_type="Sperm"
		/note="Vector: pbeloBAC11; site_1: HindIII; site_2:
		HindIII"
BASE COUNT	204 a	71 c 44 g 571 t
ORIGIN		
Query Match	5.7%; Score 79.4;	DB 17; Length 890;
Best Local Similarity	48.4%;	Pred. No. 3.8e-05;
Matches 279;	conservative 0;	Mismatches 296; Indels 2; Gaps 2
OY	AAATAAANGGGTTTTTATTAACACTATTACTATGATTCCTTCGTAAATTGTATAT	587
Dd	793 AATAGTGTTAGTATATATATAATAGATTAGCAATAATATAAAGAAGTAAAAATCGAAAAG	734
OY	588 AAATGAGACTTAATGCGAGTCGAAGAAATTCCTGGTTTAGGAGCTATTAGAGAGATTA	647
Dd	733 ATGTGAATAGAAAGAGCAAGTAATGAATAAAGTAAAAAATATATAAAAATATATATAAAAA	674

QY	648	TAATTTAACATTTAAACATGATAGTATGAATATATATATATCAATCGTTCTATTGAT	-A	706
Dd	673	TTAATATATGAATTTATATATAAAAAATTTATATTATATATATAAAAAATTA	AAAAATATTA	614
QY	707	AATTTAGCATATTTTTCGAACAGCATTTAAATCCAAAAGACGTGAAAATTTATACACAGTT		766
Dd	613	AAATTAATAATATATAAAAAAATAATTAATAAAATTTAAATTTAAATAATAATAAT		554
QY	767	ATTATATCTATAA-CCTTTTWTAAAGACACTTCGGGAAACCCCTTACGATATCATACAGA		825
Dd	553	TAAATATATATATATATATATATATATATAATAATAATAATAATTTAATATATATAA		494
QY	826	TATTATTTAATPCCAGTACGCTTCTAGTCTAAGATGTTCAATTGAAAATATTAACAGAT		885
Dd	493	TATTATTTAAT		434
QY	886	TATATGTTTTGACCAATAGCGCATCGATCTAAGCGAAATTCGAATATATATATATGA		945
Dd	433	TAAAAAATATATAATATATATAATAATAATAATAATAATAATAATAATAATAATA		374
QY	946	AGGTTATATTAATGACATAAAATTTATATATAAAAGATATACACCTAATATAGCAATAGAT		1005
Dd	373	ATTAATTTAAAAATTTAAAAAATAATAATAATAATAATAATAATAATAATAATAATA		314
QY	1006	TCCTTTGTTAAATCGGTCATTTTATTTATATATATATATATATATATATATATATATAT		1065
Dd	313	AAATTAATAAT		254
QY	1066	ATTGTAGCTTATCCGAAGATGCAATGCTTTTAAATA		1102
Dd	253	TATAAATTTAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA		217

RESULT 13

CNS0DEPO/c

LOCUS

DEFINITION

Accession

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNS0DEPO

1101 bp

DNA

linear

GSS 04-JUN-1999

Drosophila melanogaster genome survey sequence T7 end of BAC:

BACR2P01 of RPc1-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.

AI069493

GI:4949636

GSS.

Drosophila melanogaster.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque1@genoscope.cns.fr - Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES      Location/Qualifiers
source         1..1101
               /organism="Drosophila melanogaster"
               /db_xref="taxon:7227"
               /clone="BACR29P01"
```

```

/clone_11d="RPCI-98"
/note="end : T7"
BASE COUNT      289 a    155 c    118 g    307 t    232 others
ORIGIN
Query Match          5.6%, Score 78.6; DB 17; Length 1101;
Best Local Similarity 36.7%; Pred. No. 5.2e-05;
Matches 226; Conservative 97; Mismatches 283; Indels 10; Gaps 2;

OY 485 TAACCTTTTAACTTACCATGATAATTAAATGCTTATTTAGCAATAAATGGTTTTTA 544
Db 1101 TTAGRTRTAGCMTTWTATBDWTAAGRAAATTTTDTITAGCTGAATPRTGTWTWK 1042
OY 545 TAACATTACTAATGATGATTAATTCCTTGCTAAATTTGNATTAATGAGTACTTANGC 604
Db 1041 ATTTTCTTTTDMRTGTATRRRTCTTAKTKBARRWMTADGATGAGTTDGTTTTR 982
OY 605 GAAGCGCAAAATTCCTGTTTAGAGCTATTAGAGAGATATATATTAACATTAAAAC 664
Db 981 TTTTRKDKAMAMWWTTTTRKWIGTWGMMWTYAKPKRTTDDWTTTADTTAAATGAGTGA 922
OY 665 TAGATAGATGTAAATTAATATATCAATFACGTTTCATTGATAAATTTAGATATTTGCA 724
Db 921 MAKTTRMAMWMTWATFAKKFTTRTAGARTRTKTGTRRATRTDTTDKAKAGTTTGTATRGAG 862
OY 725 AAGCAT---TAATCCAAAAGAGATTGAAAAATTAATACACAAGTTATTTATCTATAACCT 781
Db 861 AKAGATTTTWTATGTAAAAATADTACGMWATFAAAATTA AAAARAAMWTTTWTATATMAWA 802
OY 782 TTTTAGAGACTTCCTGGGAAACCTTTACGATATGATACGAATATTAATTAATACAG 841
Db 801 ATTRTTAAMAAARMUDRTAAAMAAATATATATRTTAAAMAAATTTTATTAATAATTAATT 742
OY 842 TAGCTTCTAGTCTTAAGAGTGTCAATTGAAAAAATATACACAGATTATATGATTTTGCAA 901
Db 741 WAMTWTWTMTTWTGATTAAMWMAAAAAAAMKATKTBABATTAATPTFWATWTATATAT 682
OY 902 ATGCCCATCGTATCTAACGGAATTTGAATATATATTATAGAAGGTTATATTAATGCAC 961
Db 681 ATWATTTTATTTTATTAAMWTAADMATTAATTAATTAAMTAATTTATTAATAATTAAMA 622
OY 962 TAAATTTATTTATATAAA-----AGATATACACCTAATAATGAATAGATCTTTGCTT 1014
Db 621 TWAAAAAATTTTARADATTAATAATTMAAMTATAMATTAAMAAATTTWGAAAAATTTWTT 562
OY 1015 AAATCAGGTGATTTATTTAAATTTATGTATATCATATTAACAATATGAGCACATTTAGCT 1074
Db 561 ATATTAMTWTWTTTKTAAKTAAMAAATRTTAVATTAATTTATWATTTTWTWTTTWT 502
OY 1075 TATCGGAAGAGATGGA 1090
Db 501 TATTTAAATTAACAA 486

RESULT 14
CONS00396
LOCUS       CENS00396              1101 bp     DNA             linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPrT-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL063921
VERSION    AL063921.1 GI:4941778
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster.
ORGANISM  Drosophila melanogaster.
REFERENCE  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
AUTHORS    Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
TITLE       Ephydroidea; Drosophilidae; Drosophila.
JOURNAL    Genoscope.
DIRECT SUBMISSION
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 19191006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
```


us-09-816-467-3.rst

us-09-816-467-3.rst

us-09-816-467-3.rst

